

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: March 12, 2005, 13:08:19 ; Search time 887 Seconds
(without alignments)
7055.073 Million cell updates/sec

Title: US-10-627-132-30
Perfect score: 5648
Sequence: 1 MEASAGLVAGSHNRNELVLI.....IDPFIVRTKGPDPVRCGINC 1052

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5537552 seqs, 2974263231 residues
Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10627132/runat_04032005_113220_21246/app_query.fasta_1.1223
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MAIRIX=Blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MOD=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10627132 @CGN 1 1 463 @runat_04032005_113220_21246
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DISPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5648	100.0	3443	17	Sequence 29, Appl
2	5520	97.7	3589	17	Sequence 24930, A
C	3	5347.5	94.7	3448	Sequence 15093, A
	4	5055	3028	16	Sequence 29, Appl
5	4788.5	84.8	3517	9	Sequence 13, Appl
6	4561	80.8	2830	18	Sequence 139789,
7	4477	79.3	3851	17	Sequence 69, Appl
8	4395	77.8	2687	17	Sequence 833, Appl
9	4091	72.4	3776	9	Sequence 3, Appl
10	4089	72.4	3505	17	Sequence 31130, A
11	4089	72.4	3783	18	Sequence 86998, A
12	4073.5	72.1	3936	9	Sequence 9, Appl
13	4073.5	72.1	3969	16	Sequence 49, Appl
14	4073.5	72.1	3969	16	Sequence 17, Appl
15	4073.5	72.1	3969	16	Sequence 37, Appl
16	4073.5	72.1	3969	17	Sequence 49, Appl
17	4073	72.1	4297	18	Sequence 44334, A
C	18	4059	71.9	4225	Sequence 72205, A
	19	4005	70.9	3568	Sequence 41, Appl
20	4005	70.9	3568	16	Sequence 41, Appl
21	4005	70.9	3568	17	Sequence 45185, A
22	3982.5	70.5	3814	18	Sequence 7, Appl
23	3974.5	70.4	3786	9	Sequence 17, Appl
24	3974.5	70.4	3813	16	Sequence 25, Appl
25	3974.5	70.4	3813	16	Sequence 45, Appl
26	3974.5	70.4	3813	16	Sequence 45, Appl
27	3974.5	70.4	3813	17	Sequence 17, Appl
28	3974.5	70.4	3987	18	Sequence 45192, A
29	3919.5	69.4	3962	18	Sequence 72402, A
30	3917.5	69.4	3812	18	Sequence 101245,
31	3914.5	69.3	3799	16	Sequence 5, Appl
32	3914.5	69.3	3799	16	Sequence 21, Appl
33	3914.5	69.3	3799	17	Sequence 21, Appl
34	3912.5	69.3	3746	16	Sequence 29, Appl
35	3912.5	69.3	3746	16	Sequence 49, Appl
36	3912.5	69.3	3773	16	Sequence 9, Appl
37	3912.5	69.3	3773	16	Sequence 5, Appl
38	3912.5	69.3	3773	17	Sequence 9, Appl
39	3903.5	69.1	3704	16	Sequence 13, Appl
40	3903.5	69.1	3704	16	Sequence 13, Appl
41	3903.5	69.1	3704	17	Sequence 13, Appl
42	3903.5	69.1	3744	18	Sequence 100413,
43	3901	69.1	3614	15	Sequence 9, Appl
44	3899	69.0	3725	16	Sequence 45, Appl
45	3899	69.0	3725	16	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-627-132-29
; Sequence 29, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhuga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31

```
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-29

Alignment Scores:
Pred. No.: 0 Length: 3443
Score: 5648.00 Matches: 1052
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-627-132-30 (1-1052) x US-10-627-132-29 (1-3443)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 92 ATGAGGCGCCAGCGCGGCTGCTGGCCGGCTCGCACACCGAAGAGCTGGTCTGATC 151

Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 152 CCGGGCCACGAGGACCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGTGCGAGATATGC 211

Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 212 GGCACGAGGTGCGGGCTCACGGTGGACGGCGACCTCTTCGTGCGCTGCAACGAGTGGCG 271

Qy 61 PheProValCysArgProCysTyrGluTyrArgArgGluGlyThrGlnAsnCysPro 80
Db 272 TTCCTCCGTGTCGGCCCTCTACAGATACGAGCGCGGGAGGGACGCGAATCTGCCCC 331

Qy 81 GlnCysIleThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 332 CAGTGCAAGACGCGCTCAAGCGCTCAAGGGAGCGCGAGGTTGCGGGGACGATGAC 391

Qy 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db 392 GAGGAGGACATCGACGACTCGAGACACGAGTTCAACATCGACGACGAGAAATCAGCAGAG 451

Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 452 CAGCTGGAGGGCAACATCGACAGACAGCCAGATCACCGAGGCGATCTCGACGCGAGATG 511

Qy 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProlle 160
Db 512 AGCTACGGGAGGGGCCCCGACGACGCGCAGCGCAACACACCCCGCAGATCCCGCCCATC 571

Qy 161 IleThrGlySerArgSerValProValSerGlyGluPheProlleThrAsnGlyTyrGly 180
Db 572 ATCACCGGCTCCCGCTCCGCTGCGGTGAGCGGTGAGTTTCGATTACCAACGGGTATGGC 631

Qy 181 HisGlyGluValSerSerLeuHisArgGlyHisProTyrProValSerGluPro 200
Db 632 CACGGCGAGGTCTCTCTTCCTCGACAAAGCGCATCCATCCGTACCTGTGTCTGAGCCA 691

Qy 201 GlySerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArgMetAspAspTrp 220
Db 692 CGGAGTGCACAGTGGGACGAGAAAGAAAGAGTGTGAGCTGGAAGGAGGAGGAGTGCAGCTGG 751

Qy 221 LysSerLysGlnGlyIleGlyGlyValAlaAspProGluAspMetAspAlaAspVal 240
Db 752 AAGTCAAGACGAGGATCTCTCGCGCGCGCGCGGATCCCGAAGACATGGACGCCACGCTG 811

Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260
Db 812 GCATGTGAACGACGAGCGGAGGAGCGCGCTGTGAGGAAGGTGTCCATCGCGTTCGAGCAAG 871

Qy 261 ValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePheLeuArg 280
Db 872 GTGAACCGGTACCGGATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 931
```

```
Qy 281 TyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleIleCys 300
Db 932 TACCGTATCTGCAACCCCGTCCCGAGCGCATCGGGCTGTGGCTCGTCTCCATCATCTGC 991

Qy 301 GluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProlleAsp 320
Db 992 GAGATCTGGTTGCGCATCTCTGGATCTTCGACAGTTCGCCAAGTGGTTCGCCATCGAC 1051

Qy 321 ArgGluThrTyrLeuAspArgLeuSerLeuAspTyrGluArgGluGlyGluProSerLeu 340
Db 1052 CGCGAGACGTACTCGACCGCTCTCCCTCAGGTACGAGAGGAAGGAGCGCTCGCTG 1111

Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuVal 360
Db 1112 CTGTGCGCGGTGGACCTGTTCGTGAGCAGCGTGCACCGCTCAAGGAGCGCGCTGGTG 1171

Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCys 380
Db 1172 ACCGCCAACACCGTGTCTCCATCTCGCGTAGACTACCCCGTGGACAAAGGTCTCTCTGC 1231

Qy 381 TyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400
Db 1232 TAGCTCTCCGACGACGCGGTGATGCTGACGTTCAGTTCGTGTCCGAGACGCGCCGAG 1291

Qy 401 PheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420
Db 1292 TTCGCGCGCAAGTGGGTGGCTTCTGCAAGAAGTTGGGCATCGAGCCCGCGCCCGGAG 1351

Qy 421 PheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheValGln 440
Db 1352 TTCTACTTCTCGCTCAAGGTGCGACTACTCAAGGACAAAGGTGAGCCCGCTTCGTGTCAG 1411

Qy 441 GluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuVal 460
Db 1412 GAGCGCGCGCCCATGAAGAGAGAGATATGAGAGTTCAAGGTCCGATCAACGCGCTGGTG 1471

Qy 461 AlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpPro 480
Db 1472 GCCAAGGCCATGAAGGTGCGCGCAGAGGGTGGATCATGAAGGACGCGCACCGCTGGCCC 1531

Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500
Db 1532 GGGAAACAACCCCGCAGCACCCCGCGATGATCAGGTGTTCCTGGGCCACAGCGGGCGGC 1591

Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlyLysArgPro 520
Db 1592 CACGACACCGAGGGCAACGAGCTGCCCGCTCGTGTACGTCTCCCTGTGAGAACGCGCCG 1651

Qy 521 GlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaVal 540
Db 1652 GGATTCAGCACCCACAAAGAGCGCGCCATGAACGCTCTGTATTCGCTCTCCGCGGTG 1711

Qy 541 LeuThrAlaAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys 560
Db 1712 CTGACCAACGCGCATTCATGCTCAACTTGGACTGTGATCATACTACATCAACACAGCAAG 1771

Qy 561 AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr 580
Db 1772 GCCATCCGGAGGCGATGTGCTTCTCATGACCTCTAGGTCCGCGCGGAAGGTCTGCTAC 1831

Qy 581 ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn 600
Db 1832 GTTCAGTTCGCCGAGAGGTTCGACGCGCATCGCTGACGACCGCATACGCTAACACAGGAAC 1891

Qy 601 ThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrVal 620
Db 1892 ACCGTCTTCTTCGACATCAACATGAAGGGGTGACGCGCATCCCAAGGGCCCCGTGACGTC 1951

Qy 621 GlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyPro 640
Db 1952 GGGACAGGGTTCGGTTCCTCCGCGCGCAGCGCTCTACGGCTACACACCTTCCCAAGGGACCC 2011
```


QY 641 LysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLysLysArgLys 660
DB 2012 AAGAGCCCAAGATGGTACCTGCGACTGCTGCCGCGCAAGAGCGGAAA 2071
QY 661 HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGlu 680
DB 2072 CAGGCCAAGGACGGGCTGCGGAGGGCACCGCTGATATGGGAGTAGATAGCAAGGAG 2131
QY 681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
DB 2132 ATGCTCATGTCCACATGAATTCGAGAAAGCGGTTTCGGGCGAGTCCGCGCGTTCGTCAAG 2191
QY 701 SerThrLeuMetGluGluGlyValProSerSerSerSerProAlaAlaLeuLys 720
DB 2192 TCGACGCTGATGAGAAAGCGGCGTCCCTCTTCGTCGAGCCCGCGCGCTCTCTCAAG 2251
QY 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
DB 2252 GAGGCCATCATGTATCATGCTGGGCTAGAGGACAGACCGACTGGGGCTGGAGCTG 2311
QY 741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
DB 2312 GGGTGGATCTACGGTTCGATCAGGAGGACATCTGACGGGTTCAAGATGCACTGCCGC 2371
QY 761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
DB 2372 GGGTGGCGCTCGGTACTGTCATGCGAAGCGGCGCGGTTTCAGGGGTCGCGCGCATC 2431
QY 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTTPAlaLeuGlySerValGluIlePhe 800
DB 2432 AATCTATCGGACCGTCTCAACAGAGTGTCCGGTGGGCGCTGGGTTCGTGAGATCTTC 2491
QY 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
DB 2492 TTCAGCGCGCACAGCGCCCTGCTGTACGGCTACAAGACGCGCAACCTCAAGTGGCTGGAG 2551
QY 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
DB 2552 CGCTTCGGCTTACATCAACACCACTATACCCCTTTCACCTCGCTCCCGTCTCGCCCTAC 2611
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
DB 2612 TGCACCTTCCCGCGCTGCTCTCTCACGGCAAGTTTCATATGCGCTCGATTAGCAG 2671
QY 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
DB 2672 TTGCGCAGCCTCTTCTTCATCGCCTTTCATGTCCATCTTCGCGACGGCATCTCGGAG 2731
QY 881 MetArgTrpSerGlyValSerIleGluGluTTPTrpArgAsnGluGlnPheTrpValIle 900
DB 2732 ATGCGGTGGAGGGGTGAGCATCGAGAGTGGTGGAGAAACGACAGCTTCTGGGTCTATC 2791
QY 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
DB 2792 GCGCGGTGTTCGCGCATCTTTCGCGCTGTCGAGGGCGCTGCTCAAGTCTCTCGCCGG 2851
QY 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
DB 2852 ATCGACCAACCTTCAACGCTCACTTCAAGGCCACCGCGCAGCGAGGATTCGCG 2911
QY 941 GluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleIle 960
DB 2912 GAGCTTACGCTTCAAGTGGACCACTGCTCTTCATCCGCGCCACACCGCTGCTCATCAT 2971
QY 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
DB 2972 AACGTATCGGCGTCTGTCGCGCATCTTCGAGCGCATCAACACGGGTACCGATCCCTGG 3031
QY 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1000
DB 3032 GGGCGCTCTTCGCGAAGCTCTTCTTCGCTTCTCGGTTCATGCTCCACTCTACCGCTC 3091
QY 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSerIle 1020

DB 3092 CTCAAGGGGCTCATGGGCGCGCAGAAACAGGACGCCACCGTTGTGTCTATCTGGTCCATT 3151
QY 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
DB 3152 CTGTGGCTCCATCTTCTCCCTGCTCTGGGTGAGATCGACCTTTTCATCTGTCAAGACC 3211
QY 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
DB 3212 AAGGCCCGGACGTCCAGGACGTGTGGCATCAATTGC 3247

RESULT 2

US-10-425-114-24930
; Sequence 24930, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24930
; LENGTH: 3589
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-221-C10_FLI
US-10-425-114-24930

Alignment Scores:
Pred. No.: 0 Length: 3589
Score: 5520.00 Matches: 1030
Percent Similarity: 98.96% Conservative: 13
Best Local Similarity: 97.72% Mismatches: 7
Query Match: 97.73% Indels: 4
DB: 17 Gaps: 3

US-10-627-132-30 (1-1052) x US-10-425-114-24930 (1-3589)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
DB 247 ATGAGGCGCAGCGCGGCTTGTGGCGGCTCGCACAAACCGGACGAGCTGGTGTCTATC 306
QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
DB 307 CGGGGCCACAGGACCCCAACCGCTCGGCGGCTGAGCGGCGAGTGTGCGAGATATGC 366
QY 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
DB 367 GCGCAGCAGGTTCGGGCTCACGGTCGACCGGACCTTCTCGTGGCTGCAACGAGTGGCT 426
QY 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
DB 427 TTCCCGGCTGCTCCCGGCTGCTACGAGTACGAGCGCGGAGGCGCACGCAACATGCC 486
QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaAlaGlyAspAsp 100
DB 487 CAGTGAAGACCGCTACACCGCTCAAGGGGAGCGGCTCGCCGGGAGCATGATGAC 546
QY 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
DB 547 GAGGAGCATCTGACGACCTCGGAGCACCGAGTTCAACATCGACGACGAG---AAGCAGAG 603
QY 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
DB 604 CAGCTGGAGGGCAACATGCAGAACGCGCATCCCGGAGGCGCATGCTGCACGGCAAGATG 663


```
Db      2821  AGCAGTTGCGCAGGCTCTTCTTCATCGCCCTCTTCATGTCCATCTTCGCGACGGGCATC 2880
Qy      879   LeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrp 898
Db      2881  CTGGAGATGCGGTGGAGCGGGGTGAGCATCGAGGAGTGGTGAGGAGACGAGCATGTTCTGG 2940
Qy      899   ValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuValLeu 918
Db      2941  GTCATCGCGGGGTGCGCGCACCTCTTTCGCGTCTGTCAGGGCTGCTCAAGGTCTCTG 3000
Qy      919   AlaGlyIleAspThrAsnPheThrValThrSerIleAlaThrGlyAspGluAspGlu 938
Db      3001  GCCGGATCGACCACTTACCGTCACTTCAAGGCCACCGCGGAGCGTGAGCAGCAG 3060
Qy      939   PheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProGlyThrThrLeuLeu 958
Db      3061  TTGCGCGAGCTCTACGCGTTCAAGTGGACCGCTCTCATCCCGCCACCGCTGCTC 3120
Qy      959   IleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGln 978
Db      3121  ATCATCAACATCATCATCGGGTCTGTCGCGGCATCTCCGACGCCATCAACAAACGGGTACC 3180
Qy      979   SerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyr 998
Db      3181  TCGTGGGCGCGCTCTTCGCAAGCTCTTTCGCGCTTCTGGGTCTATCGTCCACCTCTAC 3240
Qy      999   ProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrp 1018
Db      3241  CGTTCCTCAAGGGCTCATGGGCGCCAGACCGACCGCCACCATGTGCTCATCTGG 3300
Qy      1019  SerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleVal 1038
Db      3301  TCCGTGCTGCTGGCTCCATCTCTCTCCCTGCTCTGGGTGAGGATCGATCCCTTTCATC 3360
Qy      1039  ArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db      3361  AGGACCAAGGGCGCCGACGTACGGCAGTGTGGCATCAACTGC 3402
```

RESULT 3

```
US-10-437-963-15093/c
; Sequence 15093, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15093
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20969C.1
US-10-437-963-15093
```

Alignment Scores:

Pred. No.:	0	Length:	3448
Score:	5347.50	Matches:	992
Percent Similarity:	96.78%	Conservative:	30
Best Local Similarity:	91.94%	Mismatches:	29
Query Match:	94.68%	Indels:	5
DB:	18	Gaps:	4

```
US-10-627-132-30 (1-1052) x US-10-437-963-15093 (1-3448)
Qy      1   MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db      3373  ATGGAGGCGAGCGCGCGGCTGTGTGCGCGGTGCGCAACCGGAACGAGCTGTGTGTGATC 3314
Qy      21   ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db      3313  CGGGGCGACGAGGAGCCCAAGCGCTGCGGGCGCTGAGCGGCGAGCTGTGCGAGATATGC 3254
Qy      41   GlyAspGluValGlyLeuThrValAspClyAspLeuPheValAlaCysAsnGluCysGly 60
Db      3253  GGCGACGAGGTGCGCGCCGACCGTTCGACGCGCACCTCTTCTGTCGCTGCTCAACGAGTGC 3194
Qy      61   PheProValCysArgProCysTyrGluArgArgGluGlyThrGlnAsnCysPro 80
Db      3193  TTCCCGGTGTGCGCGCTGTCTACGAGTACGAGCGCGCGGAGGACACCGAGACTGCTGCC 3134
Qy      81   GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaAlaGlyAspAsp 100
Db      3133  CAGTGAAGACCCGCTCAAGCGCTCAAGGGGAGCGCGAGGTGCGCGGAGCGAGGAC 3074
Qy      101  GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnGln--- 119
Db      3073  GAGGAGGACATGTGACGACCTCGAGCACGAGTTCAACATCGACGACGAGAAACGAGAGCAG 3014
Qy      120  ---ArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db      3013  CTGCAGCAGGATCAGGATGGCATGCGAAGACAGCACATCATCGAGGCGCATGCTGCACGCG 2954
Qy      139  ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
Db      2953  AAGATGAGTACGGGAGGGGCGCCGACGACGCGGACGCGCAACACGACACCCCG---CTCC 2897
Qy      159  ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
Db      2896  CCGATCATCACCGCGCTCGCTCCGCTGAGCGGGGAGTTCCCGATATCGAACACG 2837
Qy      179  TyrGlyHisGlyGluValSerSerSerLeuHisLysArgIleHisProTyrProValSer 198
Db      2836  CATGGCCATGGCGAGTCTCTCTCTCTCTGCAAGCGCATCCACCCCTACCCGCTGTCT 2777
Qy      199  GluProGlySerAlaLysTrpAspGlyLysGluValSerTrpLysGluArgMetAsp 218
Db      2776  GAGCAGGAGTGCAGAGTGGGACGAGAGAAAGAGGTGAGCTGGAAGGAGGATGAC 2717
Qy      219  AspTrpLysSerLysGlnGlyIleLeuGlyGlyAla---AspProGluAspMetAsp 237
Db      2716  GACTGGAATTCAGACGAGGCGATCGTCCGCGCGCGCGCCCGCATCCCGACGACTACGAC 2657
Qy      238  AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla 257
Db      2656  GCCGAGTCCCATCGAACGACGAGCGGCGCGCTGTCTGAGGAAGTGTGATTCGCG 2597
Qy      258  SerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhe 277
Db      2596  TCGAGCAAGGTGAACCGGTACCGGATGGTGTATCATCTCTCCGCTCTCGTGTGCTCGGCTTC 2537
Qy      278  PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSer 297
Db      2536  TTCTCTCGGTACCGCATCTCTCACCCGCTGCGCGACGCGCATCCCGCTGTGCTCACCTCC 2477
Qy      298  IleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhe 317
Db      2476  ATCATCTCGGAGATCTGGTTCGCGCTGTCTGTGGATCTCTCGACAGTGTCCCGAGTGGTAC 2417
Qy      318  ProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGlu 337
Db      2416  CCGATCGACAGGAGACCTACCTCGACCGCTCTCCCTCCGCTACGAGCGGAGGGGAG 2357
Qy      338  ProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluPro 357
Db      |||||
```

Db 2356 CCGTCGCTGCTGCGCGGTGGACCTGTTGTCAGCACGGTGGATCGCTCAAGGACGCG 2297
Qy 358 ProLeuValThrAlaAenThrValLeuSerIleLeuAlaValAspTyrProValAspLys 377
Db 2296 CCGCTGCTCACGGCCAAACCGTGTCTCCATCTCCGCGTGCAGTACCCCGTGCACAAG 2237
Qy 378 ValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGlu 397
Db 2236 GTGTCTGCTACGCTCTCCGACGCGCGGTCCATGCTCAGTTTCGAGTGCCTGTGCGAG 2177
Qy 398 ThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArg 417
Db 2176 ACGGCGAGTTCCGCCGCAAGTGGTGCCTTCTGCAAGAAGTTTCAGCATCGAGCCCGCG 2117
Qy 418 AlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThr 437
Db 2116 GCCCGGAGTTCTACTTCTCCAGAGGTGCATACCTCAAGGACAAAGTCCATCCCAAC 2057
Qy 438 PheValGlnGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsn 457
Db 2056 TTCGTCCAGAGCGCCGCGCCATGAAGAGAGTACGAGAGTTCAAGGTGAGATCAAC 1997
Qy 458 AlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThr 477
Db 1996 CGCTGTGTGGCGAAGGCGCAGAAAGTGC CGCGCGAAGGTGGATCATGAAGGACGGGACG 1937
Qy 478 ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 497
Db 1936 CCATGGCGGGGAAACAACCCCGGACACCCCGGCGCATGATCCAGGTGTTCCTCGGCCAC 1877
Qy 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 517
Db 1876 AGCGGCGCCACGACACCGAGGGCAACGAGCTCCCGCGCTCGTCTACGTTCCCGGTGAG 1817
Qy 518 LysArgProGlyPheGlnHisIleLysLysAlaGlyAlaMetAsnAlaLeuIleArgVal 537
Db 1816 AAGCGCCCGGCTTCCAGCACCAACAAGAGCGCGCGCATGAACGCCCTCATTCGTGTG 1757
Qy 538 SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn 557
Db 1756 TCGGCGGTGCTGACGAACCGCCGTTTCATGCTCAACTTGGATTGCGATCATCATCAAC 1697
Qy 558 AsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
Db 1696 AACAGCAAGGCCATCAGGAGCGCATGTGCTCTCATGGATCCGAGGTCCGACGGAAG 1637
Qy 578 ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAla 597
Db 1636 GTTTGCTACGTGACGTTCCCGCAGAGGTTTCGACGGCATCGAGTCCACGACCGGATACGCC 1577
Qy 598 AsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
Db 1576 AACCGCAACACCGTCTTTCGACATCAACATGAAGGGGCTTGATGGGATCCAGGGCGCCG 1517
Qy 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProPro 637
Db 1516 GTGTACGTCCGACAGGGTGTGTTCAGGCGCGAGGCGCTGTACGGATACACCCACCC 1457
Qy 638 LysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLys 657
Db 1456 AAGGACCAAGAGGCCCAAGATGTGACCTGCGACTGCTGCCCTTGTCTCGGAGGAAG 1397
Qy 658 LysArgLysHisAlaLysAspGlyLeuProGluGly----ThrAlaAspMetGlyValAsp 676
Db 1396 AAGCGGAAGCACGGCAAGGACCGCCCTCCCGAGGCGCGTCCGCGCGCGAGGGATGGAC 1337
Qy 677 SerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAla 696
Db 1336 ACGGACAAAGGATGCTCATGTCCGAGATGAACCTTCGAGAGCGGTTTCGGGACGTCGCG 1277
Qy 697 AlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAla 716
Db 1276 GCGTTTCGTGACGTGACGTGATGAGGAAGGCGCGTCCCGCGCTCGTCCAGCCCGCC 1217

Qy 717 AlaLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTyr 736
Db 1216 GCCTCTCTAAGAGGCCATTCATGTTCATCAGCTGCGGTACGAGGACCAAGACCGACTGG 1157
Qy 737 GlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLys 756
Db 1156 GGTCTCGAGCTGGGTGGATCTACGGGTGATCAGGGTGCATCAGCGAGGACATCTTAACGGGTCAAG 1097
Qy 757 MetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaIlePheLysGly 776
Db 1096 ATGCACCTCCGCGGTGGAGTGTATCGATCCGGAAGGCGGCTTCAAGGGG 1037
Qy 777 SerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySer 796
Db 1036 TCAGCGCCGATCAACCTATCTGACCTCTCAACAGGTGCTCCGGTGGGCGCTCGGCTCC 977
Qy 797 ValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeu 816
Db 976 GTCGAGATCTTCTTACGCCGACACGCCGCTCTCTACGGCTACAAGAACGGCAACCTC 917
Qy 817 LysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuPro 836
Db 916 AAGTGGCTCGAGCGCTTCTCTACATCAACACCACTATCACCTTCTCTCTCTCTCTCTCT 857
Qy 837 LeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro 856
Db 856 CTCTCTGCTTACGACCCCTACCGCGTCTGCTCTCTCACCGCAAGTTCATCATGCTCT 797
Qy 857 SerIleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThr 876
Db 796 CCGATTAGCACGTTTGGAGTTTGTCTTCTCATCGCTCTCATCTCTCTCTCTCTCTCTCT 737
Qy 877 GlyIleLeuGluMetArgTrpSerGlyValSerIleGluLeuTrpTrpArgAsnGluGln 896
Db 736 GGCATCTCGAGATGAGGTGGAGCGGTGAGCATCGAGAGTGGTGGAGAACGACGACG 677
Qy 897 PheTrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLys 916
Db 676 TTCTGGGTCTATCGCGCGCTGTGCGGCACCTGTTCGCGGTGGTGCAGGGCTCTCTCAAG 617
Qy 917 ValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAsp 936
Db 616 GTCTGCGCGGGATCGACCACTTCAAGTTCAGGTTCAGGCGCACCGGACGACGAGGAC 557
Qy 937 AspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThr 956
Db 556 GACGAGTTCGCGAGCTTACGCTTCAAGTGGACCACTCTCTCTCTCTCTCTCTCTCTCTCT 497
Qy 957 LeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGly 976
Db 496 CTGCTCATCTCAACATCATCGCGCTGTGCGCGCTCTCCGACGCGCATCAACACGGC 437
Qy 977 TyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHis 996
Db 436 TCCGAGCGTGGGCGCGCTCTTCGGGAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 377
Qy 997 LeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgTrpProThrValVal 1016
Db 376 CTCTACCCCTTCTCAAGGGGCTCATGGGAGGAGAACCGGACCGCCACCATTTGTTGTC 317
Qy 1017 IleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPhe 1036
Db 316 ATCTGGTCCGTGCTCTGCTCTCATCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 257
Qy 1037 IleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 256 ACCATCAAGGCGAGGGCGCTGAGCTCAGGCGAGTCCGCGCATCAACTGC 209

RESULT 4

US-10-059-29

; Sequence 29, Application US/10209059

; Publication No. US20030163838A1

GENERAL INFORMATION:
 APPLICANT: Drugga, Kanwarpal S.
 TITLE OF INVENTION: Maize Cellulose Synthases and Uses
 TITLE OF INVENTION: Thereof
 FILE REFERENCE: 0864R2
 CURRENT APPLICATION NUMBER: US/10/209,059
 CURRENT FILING DATE: 2002-07-31
 PRIOR APPLICATION NUMBER: 60/096,822
 PRIOR FILING DATE: 1998-08-17
 PRIOR APPLICATION NUMBER: 09/371,383
 PRIOR FILING DATE: 1999-08-06
 PRIOR APPLICATION NUMBER: 09/550,483
 PRIOR FILING DATE: 2000-04-14
 NUMBER OF SEQ ID NOS: 52
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 29

LENGTH: 3028
 TYPE: DNA
 ORGANISM: Zea mays
 US-10-209-059-29

Alignment Scores:

Pred. No.: 0 Length: 3028
 Score: 5055.00 Matches: 944
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 89.50% Indels: 0
 DB: 16 Gaps: 0

US-10-627-132-30 (1-1052) x US-10-209-059-29 (1-3028)

QY 109 HisGluPheAsnIleAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsn 128
 Db 1 CACGAGTTCAACATCGACGACGAGAGATCAGCAGAGCGAGCTGGAGGGCAACATGCAGAAC 60
 QY 129 SerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAsp 148
 Db 61 AGCCAGATCACCGAGGCGATGTGTCACGCGCAGGATGAGCTACGGGAGGGGGCCCGCAGCAG 120
 QY 149 GlyAspGlyAsnAsnThrProGlnIleThrProIleIleThrGlySerArgSerValPro 168
 Db 121 GCGACGCGCAACACACCCCGCAGATCCCGCCCATCATCACCGCTCCCGTCCGTCGCG 180
 QY 169 ValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeu 188
 Db 181 GTGACGGGTGAGTTTCCGATTACCAACGGGTATGCGCCACGGCGAGGTCTCGTCTTCCCTG 240
 QY 189 HisIysArgGlyIleHisProTyrProValSerGluProGlySerAlaLysTrpAspGluLys 208
 Db 241 CACAAGCGCATCCATCCGTACCCCTGTGTGAGCCAGGAGGTGCCAAGTGGGACGAGAAAG 300
 QY 209 LysGluValSerTrpLysGluArgMetAspAspTrpLysSerLysGlnGlyIleLeuGly 228
 Db 301 AAAGAGGTGAGCTGGAAGAGAGGATGACGACTGGAAGTCCACAGGGGATCTCCGCG 360
 QY 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
 Db 361 GCGCGCGCGATCCCGAGACATGACCGCGAGCGTGGCACTGAACGACGAGCGGAGGCGAG 420
 QY 249 ProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIle 268
 Db 421 CCGCTGTGAGGAGGTGTGATCTCGATCCGTCGTCGAGCAAGGTGAACCCGTACCGATGGTGATC 480
 QY 269 ValValArgLeuValValIleAlaPhePheLeuArgTyrArgIleLeuHisProValPro 288
 Db 481 GTGGGCGGTCTCGTGTGCTCGCTCTTCCTCCCGGTACCGGTATCCTGCACCCCGTCCCG 540
 QY 289 AspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrp 308
 Db 541 GACGCCATCGGGCTGTGCTCTCCATCATCTCGAGATCTGTTTCCCATCTCTCTCTGG 600
 QY 309 IleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeu 328

Db 601 ATCTCTCGACCAAGTTCCCAAGTGGTTCCTCCATCGACCGCGAGAGCTACTCTCGACCGCTC 660
 QY 329 SerLeuArgTyrGluArgGluGlyProSerLeuLeuSerAlaValAspLeuPheVal 348
 Db 661 TCCCTCAGTACGAGAGGAGGAGCGCTGCTGCTGCTGCGCGGTGAGCTGTTCGTG 720
 QY 349 SerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIle 368
 Db 721 AGCAGCGTGGACCGCTCAAGGAGCGCGCTGGTGTGAGCGGCAACACCGTGTCTTCCATC 780
 QY 369 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 388
 Db 781 CTCGCGGTAGACTACCCCGTGGACAGGTCTCTCTGCTACGTCTTCGAGCAGCGCGCGTGC 840
 QY 389 MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPhe 408
 Db 841 ATGCTGAGTTCGAGTCTGCTCGGAGACGCGAGTTCGCGCGCAAGTGGTGGTCCCTTC 900
 QY 409 CysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAsp 428
 Db 901 TGCAGAAGTTCGCGCATCGAGCCCGCGCGCGGAGTTCCTACTTCTCGCTCAAGGTCGAC 960
 QY 429 TyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGlu 448
 Db 961 TACCTCAAGGACAAAGGTGACGCCCTTCGTGAGGAGCGCGCGCATGAAGAGAGAG 1020
 QY 449 TyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAla 468
 Db 1021 TATGAGGATTCAGGTCCGATCAACCGCTGTGTGCGCAAGGCCCATGAAGTGGCGGCA 1080
 QY 469 GluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 488
 Db 1081 GAGGGTGGATCATAGGACGCGACGCGGTGGCGCGGGAACAACACCGCGGACCCACCC 1140
 QY 489 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeu 508
 Db 1141 GGCATGATCCAGGTGTCTCTGGGCGCACAGCGCGCGCGCACACGAGGGCAACAGGCTG 1200
 QY 509 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAla 528
 Db 1201 CCGCGCTCTGTGATCTCCGTGAGNAGCGCGCGGATTCAGGACCCACACAGAGGCC 1260
 QY 529 GlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeu 548
 Db 1261 GCGCCATGAACGCTCTGATTCGGGTCTCCCGCTGTGTCGACCAACCGCGCATTCATGCTC 1320
 QY 549 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhe 568
 Db 1321 AACTTGGACTGTGATCACTACATCAACAACAGCAAGGCCCATCCGGGAGGCCATGTGCTTC 1380
 QY 569 LeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAsp 588
 Db 1381 CTCTATGACCTCTAGTTCGCGCGGAGGTCTGCTACTGTTTCCGCGAGAGGTTCCGAC 1440
 QY 589 GlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMet 608
 Db 1441 GGCATCGAGTGCACGACCGCATCGCTAACAGGAACAACCGTCTTCTTCGACATCAACATG 1500
 QY 609 LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg 628
 Db 1501 AAGGGCTGGACGCGCATCAAGGGCGGTGTACGTGCGGACAGGGTGTGCTTCCGGCGC 1560
 QY 629 GlnAlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThrCys 648
 Db 1561 CAGCGCTCTACGGCTACACCTCCCAAGGAGCCCAAGAGGCCCAAGTGTGTACCTGC 1620
 QY 649 AspCysCysProCysPheGlyArgLysLysArgLysHisAlaLysAspGlyLeuProGlu 668
 Db 1621 GACTGCTGCCGTGCTTCCGCGCGCAAGAGCGGAAACACGCGCAAGGACGGGTGCGGAG 1680
 QY 669 GlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPhe 688

Qy 897 eTTPValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysVa 917
Db 2816 TTGGGTCAATTGGTGTATCAGCTCACCTCTTTGCTGTTATATACAAGGTCTGCTAAAGGT 2875
Qy 917 lLeuAlaGlyLeuAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAs 937
Db 2876 TCTGGCTGGAATTGACACCAATTTCACTGTTTACATCAAGGCAACA-----GATGATGA 2929
Qy 937 pGluPheAlaGluLeuValAlaPheLysValThrThrLeuLeuIleProProThrThrLe 957
Db 2930 AGAGTTTGGGAATTGACACCTTTAAGTGACATACACTCTTGATTTCTCAACACCATAT 2989
Qy 957 uLeuIleAlaValIleGlyValAlaAlaGlyLeuSerAspAlaIleAsnGlyTy 977
Db 2990 TTTGATCATTAACATTGTTGGTGTGTTGCTGGAATCTCAGATGCCATAAACAATGGTA 3049
Qy 977 rGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLe 997
Db 3050 CCAATCTCGGGGACCACTCTTTGGAAAGCTCTTCTTTCTCTGCGGTGATTTGCCATCT 3109
Qy 997 uTyrProPheLeuLysGlyLeuMetGlyValGlnAsnArgThrProThrValValIle 1017
Db 3110 CTATCCATTCTTAAAGGTTTGATGGGTGCGCAAAATCGCACCCACCATTTGTTGGAT 3169
Qy 1017 eTrpSerIleLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPropheI 1037
Db 3170 TTGGTCAGTGCTATTTGGCTCTATTTCTCTTACTTTGGTAAAGATTGATCCATTTGT 3229
Qy 1037 eValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3230 CCTCAAGACTAAGGACCTGATACCAAGCTATGTGAATCAACTGC 3275

RESULT 6

US-10-425-115-139789
; Sequence 139789, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 139789
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_58974C.1
US-10-425-115-139789

Alignment Scores:
Pred. No.: 0 Length: 2830
Score: 4561.00 Matches: 856
Percent Similarity: 99.08% Conservative: 8
Best Local Similarity: 98.17% Mismatches: 6
Query Match: 80.75% Indels: 3
Db: 18 Gaps: 1

US-10-627-132-30 (1-1052) x US-10-425-115-139789 (1-2830)

Qy 183 GluValSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySer 202
Db 2 GAGCTCTCGTCTTCCCTGCACAAAGCGCATCCATCCCTACCTGCTGCTGAGCGAGGAGC 61
Qy 203 AlalysTrpAspGluLysGlyValSerTrpLysGluArgMetAspAspTyrLysSer 222
Db 62 GCCAAGTGGGACGAGAAAGAGAGTG-AGCTGAAGGAGAGGATGACGACTGGGAAGTCC 120

Qy 223 LysGlnGlyIleLeu-----GlyGlyGlyAlaAspProGluAspMetAspAlaAspVal 240
Db 121 AAGAGGGGCATCTCCGCGCGCGCGCGCGCGACACCCCGAAGACATGAGCGCGCAGCTG 180
Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260
Db 181 CCCTGAACACACGAGCGAGGCGCGCTGTCGCGAAGGTGCGATCGCGTCAAGCAG 240
Qy 261 ValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePheLeuArg 280
Db 241 GTGACCCCGTACCGGATGGTCACTCGTGGTGGCTCTCGTCTCTCGCTCTTCTTCTCCGC 300
Qy 281 TyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleLeuCys 300
Db 301 TACCGCATCTCGACCTGCTCCGCGACGCCATCGGGCTGTGGCTCGTCTCATCATCTGC 360
Qy 301 GluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsp 320
Db 361 GAGATCTGGTTCGCGCTGCTCGATCTCTGACCTGAGTTCCTCCCAAGTGGTTCCCATCGAC 420
Qy 321 ArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeu 340
Db 421 CGCGAGAGTACTTGGACCCCTCACTCGGTACGAGAGGAGGAGGAGCGCTCGCTG 480
Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuVal 360
Db 481 CTGTGCGCGGTGGACCTGTTCTGTCAGACGCGTGGACCGCTCAAGGAGCGCGCTCGTGC 540
Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCys 380
Db 541 ACCGCCAACACCGTGTCTCCATCTCTCGCTGGACTACCCCGTGGACAAGGTCTCTGTC 600
Qy 381 TyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400
Db 601 TAGTCTCCGACGACGCGCGCTCGATGCTCAGCTTCGAGCGCTGTGCGAGACGCGCGAG 660
Qy 401 PheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420
Db 661 TTGCGCGGGAAGTGGGTGCGCTCTGCAAGAAGTCTGTCATCGAGCCCGCGCGCGCGAG 720
Qy 421 PheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheValGln 440
Db 721 TTCTACTTCTCGCTCAAGGTGCTTACCTCAAGGACAAGGTGCGAGCCACCTTCGTGCGAG 780
Qy 441 GluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuVal 460
Db 781 GAGCGCGCGCCATGAAGAGAGAGTACGAGAGTTCAGAGTCCGATCAACGCGCTGGTG 840
Qy 461 AlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpPro 480
Db 841 GCCAAGGCCATGAAGGTGCGCGCAGAGGGGTGGATCATGAAGGACGCGCACGCCGTGGCCC 900
Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500
Db 901 GGGAAACAACACCGCGCACCCCTCGGCATGATTCAGAGTTCCTGGGCGCACAGCGCGGCG 960
Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgPro 520
Db 961 CACGACACCGAGGCGAAGCGAGCTGCCCGCGCTCGTGTACGCTCCCGTGAGAGCGCGCG 1020
Qy 521 GlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaVal 540
Db 1021 GGCTTCCAGCACCAACAAGGCGCGCCCATGAACGCTCTGATCCGCGCTCTCGCGCGCTG 1080
Qy 541 LeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys 560
Db 1081 CTGACCAACGCGCGCTTCACTGCTCACTGACTGTGATCTACTACTACTCAACAAACAGCAG 1140
Qy 561 AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr 580
Db 1141 GCCATCCGGAGGCGCATGTGCTTCTCATGACCCCTCAGGTTCGCGCGGAGGTCTGCTAC 1200
Qy 581 ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn 600

Db 333 TTTCTGTCTGTGGCGTGTCTAGTACGAGACGAGAAATCAAGTCGTGCCG 392
Qy 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 393 CAGTGCATACTCTTACAGCGTCAAAAAGGGAGTCCACGGGTGGAGGTGACGATGAT 452
Qy 101 GluGluAspIleAspAspLeuGluHisGluPheAenIleAspAspGluAenGlnArg 120
Db 453 GAAGAAGCGTGTATGACATAGACATGATTTATGTGGAGATCTACCAAGAAACAGG 512
Qy 121 GlnLeuGluGlyAenMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 513 CAG-----CAGATCACCGAGGCGATGCTCCACGACGCGCATG 548
Qy 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAen-----AenThrProGlnIlePro 158
Db 549 AGCTATGCCAGGTCCGACGACGAGAAATTCGCAGATGTCTCATATCCAGAGCTTCCT 608
Qy 159 -----ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIle 175
Db 609 CCGCAGATTCCTGTACTTGCAACCGGCCACTCGGT---GTGAGTGGGAGATTCGA--- 662
Qy 176 ThrAenGlyTyrGlyHisGlyGluVal-----SerSerSerLeuHisLysArgIleHis 193
Db 663 ACGTCACTACTACGCGACACAACTTGTCTGCCAACCTCGCAATGCTGAAGCGTGTGCAT 722
Qy 194 ProTyrProValSerGluProGlySerAlaLys-----TrpAspGluLysGlu--- 210
Db 723 -----CCAAGCTCCGAGCCGGGAGTGAAGGATCATCATGATCCAAACAGGGATATT 776
Qy 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLys 223
Db 777 GGTCTTATGCTTTGGGAACTGTCTCTGGAGGCGAGCGATGGTTATTAATCGAAG 836
Qy 224 GlnGlyIleLeuGly-----GlyGly--- 230
Db 837 GAAACAAATCAGGCCCATGTCATGATACGGAAGGAGATATCAATATAATGGGGGT 896
Qy 231 -----AlaAspProGluAsp---MetAspAlaAspValAlaLeuAenAspGluAlaArg 247
Db 897 GCACAAATGAGCCCTGAAGATTATATGATCCCGATATGCCAATGACCGATGAAGCAAGG 956
Qy 248 GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetVal 267
Db 957 CAGCCACTGTCCGAAAAGTGCATTCCTTCACGAAATAAATCCATACCGAATGCTC 1016
Qy 268 IleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProVal 287
Db 1017 ATTGTAATTCGACTGATAGTGTGGGTATTTTCTCCGCTATCGTCTCTGAAATCCAGTG 1076
Qy 288 ProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSer 307
Db 1077 AAGAAATGATATGCGGCTCTGGGCCACTTCTATCGTGTGTGAATCTGGTTTGGCTTGTCA 1136
Qy 308 TrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrIleuAspArg 327
Db 1137 TGGATTCTTGATCAGTTTCCCAAGTGGTTCCTATCAGTCGTGGAACGTAICTTGATCGA 1196
Qy 328 LeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe 347
Db 1197 CTGTCAATTAAGGTACGACGAGAGGCGAACCATCAATGCTTGCACCTGTTGACCTCTTT 1256
Qy 348 ValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSer 367
Db 1257 GTGAGTACTGTAGATCCACTGAAGGAGCGCTCTTTGGTTACTTGCCAAATACAGTATTATCA 1316
Qy 368 IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAla 387
Db 1317 ATCCCTTCAGTAGACTACCTCTAGCAATGTCCTCTGTATGTCTCTGATGACGAGCG 1376
Qy 388 SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro 407
Db 1377 TCGATGCTTACTTTTGAATCTCTCTGAGACCTCAGAAATTTTGCAGAAAATGGGTACCA 1436

Qy 408 PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysVal 427
Db 1437 TTCTGCAAGAAATTCGACATTTGAGCCTCGCGCTCCCGAAATCTATTCTCTCGAATAATT 1496
Qy 428 AspTyrIleuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg 447
Db 1497 GACTATCTGAAGGACAAATTTCAACCCACCTTTGTCAAAGAGCGCGCTGCATGAAGAGA 1556
Qy 448 GluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLeuMetLysValPro 467
Db 1557 GAATATGAAGAATTCAGGTGCGCATCAATCGGTGGTTGGTTCGAAAGGCGCTCTAAAGTGCC 1616
Qy 468 AlaGluGlyTrpIleMetLysAspGlyThrProTrpProGlyAenAenThrArgAspHis 487
Db 1617 AAGGAAGGATGGCAATGCAAGCGGTACGCTTGGCTTGGTAAATAATACCGTGACCAT 1676
Qy 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGluAsnGlu 507
Db 1677 CTTGGTATGATCCAAAGTGTCTTTGGGTACAGTGGCGGCTCGATACAGAGGCAATGAG 1736
Qy 508 LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLys 527
Db 1737 CTTCTCGGCTAGTATATGTTTCTCGTGAAGAGAGCTGTTTCCAGCATTCACAGAGAG 1796
Qy 528 AlaGlyAlaMetAenAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMet 547
Db 1797 GCCGTGCCATGAATGCTTTGGTTCGGGTTCCTGCTGCTCACCANTGCTCCATTATG 1856
Qy 548 LeuAenLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
Db 1857 CTGAATCTGGATTGTGATCACTACATTAACAATAGCAAGCAATCAGAGGAGCATGTGC 1916
Qy 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGluArgPhe 587
Db 1917 TTTATGATGATCCTCAGGTGGGAGAAAGTCTGTTATGTCCCAATTCCTCCAGAGATT 1976
Qy 588 AspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsn 607
Db 1977 GATGATATGATCCCAATGACCGTTACGCCAATCGAACACCGTATTCTTTGATATCAAC 2036
Qy 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArg 627
Db 2037 ATGAAAGTCTCGATGGAATTCAGAGGCTGTATATGTGGAACTGGATGCTGATTCAGA 2096
Qy 628 ArgGlnAlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThr 647
Db 2097 AGACAAGCTCTATATGGGTATGGGCTCCCAAGGCCCAAAACGTCCTCCAGATGCTGACC 2156
Qy 648 CysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLys----- 663
Db 2157 TGTGATTGCTCCCTTGTGCGGTCTCGTAAAGAGTCTCCGAAAGAAATAGTAGCAAG 2216
Qy 664 -----AspGlyLeuProGluGlyThrAla 671
Db 2217 AAAAGTCAGGAATCCCGCTCCCGCTACCAATCTGGACGGGATCGAGAGAGTAGAA 2276
Qy 672 AspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAenPheGluLysArg 691
Db 2277 -----GGTTATGATGACGAAAGAGCATTTGTTGATGACCAACTAGACTTCGAGAGAGAG 2330
Qy 692 PheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProPro 711
Db 2331 TTTGGCCAGTCTTCAGCTTTTGTTCATCCACTCTGATGGAGAAATGGTGGTTCGCGAA 2390
Qy 712 SerSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGlu 731
Db 2391 ACAGCAATTCAGCTGAATTTGTAAGGAGGCTATTTCATGTATCAGCTGTGGATATGAA 2450
Qy 732 AspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIle 751
Db 2451 GACAAAACGGAATGGGAAAAGAGCTTGATGGATCTATGGATCAGTCACAGAGACATT 2510

QY	752	LeuThrGlyPheIysMetHisCysArgGlyTTPArgSerValTyrCysMetProLysArg	771
DB	2511	CTGACTGGATTCAAGATGTCACACTCGAGGCTGGCGGTCCATTACTGTATGCCCAACGA	2570
QY	772	AlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAenGlnValLeuArg	791
DB	2571	GCAGCAATCAAGGGTCTGCTCCNATCAATCTATCAGACCGTTTGAACCAAGGTGTGCGT	2630
QY	792	TrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyr	811
DB	2631	TGGGCTTTGGGATCAGTAGAAATTTTTCATGAGCAGACATTGCCCAATCTGGTATGGCTAT	2690
QY	812	LysAsnGlyAsnLeuLysTyrLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro	831
DB	2691	---GGGGGAGGCTCGAAATGGCTTGAAGAAGTTTGGCTATATCAACCACTGTGCTATCCA	2747
QY	832	PheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGly	851
DB	2748	TTCACTCTCTTCCACTCATTTGGCTATTGGACACTTCCAGCGGTCAGTTTGTCTCACTGGC	2807
QY	852	LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMet	871
DB	2808	AAATTTGTGATCCTTCAGACTCAGTACTTTTGGCAAGTCTATTTTTTAATAGCTCTTTTCAATC	2867
QY	872	SerIlePheAlaThrGlyIleLeuGluMetArgTTPSerGlyValSerIleGluGluTTP	891
DB	2868	TCAATTTTTCCTCACTGGTATTTCTGAAATCAGGTGGAGTGGAGTGGAGCATTTGAAGAATGG	2927
QY	892	TrpArgAsnGluGlnPheTTPValIleGlyGlyValSerAlaHisLeuPheAlaVal	911
DB	2928	TGGCGAATGAACAGTTCCTGGGTATTGGAGGGGTTCCTGCACATTTTTTTTGAGCATTA	2987
QY	912	GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla	931
DB	2988	CAAGGCTGCTCAAGGTACTGGCAGGCATTGTATACAAATTTTCACAGTCACTGCCAAGGCA	3047
QY	932	ThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTTPThrThrLeuLeu	951
DB	3048	TCA-----GATGACGGTGAATTTGGGGAACTGTGTGCAVTCAAATGGACCACTCCTC	3101
QY	952	IleProProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAsp	971
DB	3102	ATTCTCTCTACAAACCTGCTGTGTCATACACCTTGTGGGGGTGGTGTGGCGTAGCAGAT	3161
QY	972	AlaIleAsnAsnGlyTyrGlnSerTTPGlyProLeuPheGlyLysLeuPhePheAlaPhe	991
DB	3162	GCAATCAACAATCGAATTTCACTAGTCATCGGGTCTCTCTTGGGTAAAGCTTTTCTTTGCAATC	3221
QY	992	TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr	1011
DB	3222	TGGGTCAATTTGTGACCTGTATCTCTTCTCAAGGGTCTCATATGGCGCAGGACCAACCA	3281
QY	1012	ProThrValValValIleTTPSerIleLeuLeuAlaSerIlePheSerLeuLeuTTPVal	1031
DB	3282	CCACCAATCTGGGTATTATTGGTCAATTCCTGGCATCTGTTTCTCTCTTTTCTGGGTA	3341
QY	1032	ArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn	1051
DB	3342	AGAATTGATCTTTCTTGAGTAAGGTTAAAGCCCGACAGTACTAAACAATGTGCATCAAC	3401
QY	1052	Cys	1052
DB	3402	TGC	3404

RESULT 8

US-10-425-114-833

US-10-423-114-833
; Sequence 833, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

```

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 833
; LENGTH: 2687
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700093978_FLI
US-10-425-114-833

```

Alignment Scores:

Alignment Scores:					
Pred. No.:	0	Length:	2687		
Score:	4395.00	Matches:	824		
Percent Similarity:	99.88%	Conservative:	0		
Best Local Similarity:	99.88%	Mismatches:	1		
Query Match:	77.82%	Indels:	0		
DB:	17	Gaps:	0		

US-10-627-132-30 (1-1052) X US-10-425-114-833 (1-2687)

228	Qy	GlycylglylalaaspProGluaspMetaspAlaaspValalaLeuasnAspGlualaArg	247
1	Db	GGCGGGCGCCGATCCCGAAGACATGGACGCCGACGTGGCACGTGAACGACGAGCGGAGG	60
248	Qy	GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetVal	267
61	Db	CAGCCGCTGTCGAGGAAGGTGTGCATCGCGTCGAGCAAGGTGAACCCGTACCGGATGGTG	120
268	Qy	IleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProVal	287
121	Db	ATCGTGGTGGCTCTCGTTGTGCTCGCTTCTTCCTCCGGTACCGGTATCTGCACCCCGTTC	180
288	Qy	ProAspAlaIleGlyLeuTyrLeuValSerIleIleCysGluIleTyrPheAlaIleSer	307
181	Db	CCGAGCGCCATCGGGCTGTGGCTCGTCTCCATCATCTCGAGATCTGGTTCGCCATCTCC	240
308	Qy	TyrIleLeuAspGlnPheProLysTyrPheProIleAspArgGluThrTyrIleuAspArg	327
241	Db	TGGATCTCTCACCAGATTCCCAAGTGGTTCGCCATCGACCCGAGACGCTACCTCGACCGC	300
328	Qy	LeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe	347
301	Db	CTCTCCCTCAGGTACGAGAGGAAGGGAGCGCGTGCCTGCTGCGCGGTGACCTGTTC	360
348	Qy	ValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSer	367
361	Db	GTGAGCAGGTGGACCCGCTCAAGAGCGCGCGTGGTGGACCGCCAAACCGTGTCTCTCC	420
368	Qy	IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAla	387
421	Db	ATCCTCGCCGTAGACTACCCCGTGGAACAAGGTCTCTGCTACGTCTCCGACGACGGCGCG	480
388	Qy	SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro	407
481	Db	TCGATGTGACGTTTCGAGTGCCTGTTCGGAGACGGCCGAGTTTCGCGCGCAAGTGGGTGCC	540
408	Qy	PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysVal	427
541	Db	TTCTGCAAGAAGTTTCGGCATCGAGCCCGCGCCCGGAGTTTCTACTTCTCGCTCAAGGTTC	600
428	Qy	AspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg	447
601	Db	GACTACTCTCAAGGACAAGGTTCGACCCACCTTCTGTCGAGGAGCGCCGCCATGAAGAGA	660
448	Qy	GluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValPro	467
661	Db	GAGTATCAGAGATTCAAGGTCGAGTCAACCGCTGGTGGCCAAGGCCATGAAGTGGCG	720

QY 468 AlaGluGlyTrpIleMetLysAspGlyThrProTrpProGlyAenAenThrArgAspHis 487
DB 721 GCAGAGGGGTGGATCATAGAGGACGGCACCGGTGGCCCGGGAACAACACCCCGACAC 780
QY 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAenGlu 507
DB 781 CCGGCGCATGATCCAGGTGTTCTTGGGCGCACAGCGGGCGCCACACACCGAGGGCAAGAG 840
QY 508 LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLys 527
DB 841 CTGCCCCGCTCGTGTACGTCTCCGTGAGAGCGCCCGGGATTTAGCACCAACAAG 900
QY 528 AlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAenAlaProPheMet 547
DB 901 GCCGGCGCCATGAACGCTGTGATTCGGGTCTCGCGGTCTGACCAACAGCGCCATTTCATG 960
QY 548 LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
DB 961 CTCACTTGGACTGTGATCTACTACATCAACACAGCAAGGCCATCCGGAGGCCATGTGC 1020
QY 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPhe 587
DB 1021 TTCTCTATGACCTCAGGTCCGGCGGAGGCTCTGCTACGTTTCAGTTCCCGCAGAGGTTTC 1080
QY 588 AspGlyIleAspValHisAspArgTyrAlaAsnArgAenThrValPhePheAspIleAsn 607
DB 1081 GACGGCATCGACGTCGACGCCGATACGCTAACAGGAACACCCGTCTTCTTCGACATCAAC 1140
QY 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArg 627
DB 1141 ATGAAGGGGTGGACGGCATCCAGGCCCGGTGTACGTCCGGGACAGGGTTCGGTTCGGG 1200
QY 628 ArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThr 647
DB 1201 CGCCAGGCGCTCTACGGCTACAACCCCTCCCAAGGAGCCCAAGAGGCGCCCAAGATGGTGACC 1260
QY 648 CysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLysAspGlyLeuPro 667
DB 1261 TGCAGACTGCTCCCGTCTCGCGCGCAAGAGCGGAAACACGCGCAAGGACGGGCTCGCG 1320
QY 668 GluGlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsn 687
DB 1321 GAGGGCACCGCTGATATGGGAGTAGATAGCAGCAGGAGATGCTCATGTCTCCACATGAAC 1380
QY 688 PheGluLysArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGly 707
DB 1381 TTCGAGAGCGGTTCCGGGCGAGTCCCGCGCGTTCGTACGTCGACGCTGATGGAGGAAGGC 1440
QY 708 GlyValProProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSer 727
DB 1441 GCGGTCCCTCTCGTCGAGCGCCCGCGCTCTCAAGAGGCGCCATCATGTATCATCAGC 1500
QY 728 CysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIle 747
DB 1501 TCGGCTACGAGGACACAGACCTGGGGGCTGGAGCTGGGGTGATCTACGGGTCTGATC 1560
QY 748 ThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyrCys 767
DB 1561 ACGGAGGACATCTCAGCGGGTTCAAGATGCACTGCGCGGGGTGGCGCTCGGTGTACTGC 1620
QY 768 MetProLysArgAlaAlaPheLysGlySerAlaProIleAenLeuSerAspArgLeuAen 787
DB 1621 ATGCCGAAAGCGGGCGCGCTTCAAGGGGTTCGGCGCGATCAATCTATCGAGCGCTCAAC 1680
QY 788 GlnValLeuArgTrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeu 807
DB 1681 CAGGTGCTCCGGTGGGCGTGGGGTCCGTCGAGATCTTCTTCAGCGGCGACAGCCCGCTG 1740
QY 808 LeuTyrGlyTyrLysAsnGlyAenLeuLysTrpLeuGluArgPheAlaTyrIleAenThr 827
DB 1741 CTGTACGGCTACAAGAAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCGCTACATCAACACC 1800

QY 828 ThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCys 847
DB 1801 ACCATCTACCCCTTACCTTCGCTCCGCTGCTGCTACTGCAACCTCCCGCGCTGTC 1860
QY 848 LeuLeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIle 867
DB 1861 CTCTCTACCGGCAAGTTTCATTCGCGTTCGATTCGATTCGATTCGATTCGATTCGATTC 1920
QY 868 AlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSer 887
DB 1921 GCCCTCTTCATGTCATCTTCGCGACGGGATCTCTGAGATGCGGTGGAGCGGGTGAGC 1980
QY 888 IleGluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeu 907
DB 1981 ATCGAGGAGTGGTGGAGGAACGAGCAGTTCTGGGTTCATCGCGCGGTGTCGCGCATCTC 2040
QY 908 PheAlaValValGlnGlyLeuLysValLeuAlaGlyIleAspThrAenPheThrVal 927
DB 2041 TTCCGCGTCTGTCAGGCGCTCTCAAGGTCTCTCGCGGGATCGACACCACTTCACCGTC 2100
QY 928 ThrSerLysAlaThrGlyAspGluAspGluPheAlaGluLeuTyrAlaPheLysTrp 947
DB 2101 ACCTCAAGGCGCACCGCGACGAGGACGAGTTCCGCCGAGCTCTACGCTTCNAAGTGG 2160
QY 948 ThrThrLeuLeuIleProProThrThrLeuLeuIleIleAsnValIleGlyValValAla 967
DB 2161 ACCAGCTCTCTATCCCGCCACCGCTCTCATATTAACGTCTATCGCGCTCGTGGCC 2220
QY 968 GlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeu 987
DB 2221 GGCATCTCTCCAGCCCATCAACAACGGGTACAGTCTCTGGGGGCGCTCTTCGGCAAGCTC 2280
QY 988 PhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArg 1007
DB 2281 TTCTTCGCTCTGGGTTCATGCTCCACCTCTACCGCTCTCAAGGGGCTCATGGGGCGC 2340
QY 1008 GlnAsnArgThrProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSer 1027
DB 2341 CAGACAGGACGCCACCGTTGTTGTCATCTGGTCCATCTGTGCGCTCCATCTTCTTC 2400
QY 1028 LeuLeuTrpValArgIleAspProPheIleValArgThrLysGlyProAspValArgGln 1047
DB 2401 CTGCTCTGGGTTCAGGATCGACCTTTCATCGTCAGGACCAAGGGCGCGGAGCTCAGGCAG 2460
QY 1048 CysGlyIleAsnCys 1052
DB 2461 TGTGGCATCAATTGC 2475

RESULT 9
US-09-900-237-3
; Sequence 3, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 3776
; TYPE: DNA
; ORGANISM: Zea mays
US-09-900-237-3
Alignment Scores:


```
Db 2147 AACTGCTGGCCAAAGTGGTGCATTTGCTGTGCTGTTTGGTAACAGACCAAGAAG 2206
Qy 661 HIAALALYSASPGLYLEUPROGLU-||||| ||||||| ||||||| ||||||| |||||||
Db 2207 AAGACCAAGACCTCTAAACCTAAATTTGAGAGATAAGAAACCTTTTAAAGAAAAGGAA 2266
Qy 669 -----GlyThrAlaaspMetGlyValaspSer 677
Db 2267 AATCAAGCCCTGCATATGCTTTGGTGAATTTGATGAAGCCGCTCCAGGAGCTGAAAT 2326
Qy 678 AsplysGluMetSerHisMetAsnPheGluysArgPheGlyGlnSerAlaAla 697
Db 2327 GAAGAAGCTAGTATTGTAATCAACAGAGTTGAAAAGAAATTTGGCCAGTCTTCAGTT 2386
Qy 698 PheValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAla 717
Db 2387 TTTGTTGCATCCACATCTTTGAGAAATGTTGAACCTCGGAAGAGTGCAGTCCAGCTTCT 2446
Qy 718 LeuLeuysGluAlaIleHisValIleSerCysGlyTyrGluAaplysThrAapTTPGly 737
Db 2447 CTTCTGAAGGAGCTATACATGTCATCAGTTGTGATATGAAGACAAACAGGCTGGGA 2506
Qy 738 LeuGluLeuGlyTTPleTyrGlySerIleThrGluAspileLeuThrGlyPheLysMet 757
Db 2507 AAAGATATTGGTTGGATTATGGATCATGTCACAGAGATATCTTACTGGGTTAAGATG 2566
Qy 758 HisCysArgGlyTTPArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777
Db 2567 CACTGCCATGTTGGCGTCAATTTACTGCATACCTAAACGGCGCGCTTCNAAGGTTCC 2626
Qy 778 AlaProlleAsnLeuSerAspArgLeuAsnGlnValLeuArgTTPAlaLeuGlySerVal 797
Db 2627 GCACCTCTCAATCTTCGATCGCTTACCAGAGTCTTCGGTGGGCTCTTGGTTCATATT 2686
Qy 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817
Db 2687 GAATTTTCTTCAGCAACCATCGCCCTCTCGGTATGGGTAT---GGTGGTGGACTAAAG 2743
Qy 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837
Db 2744 TTCCTGGAAGAGTTTTCGTACATTAACCTCATCGTATACCTTGGACATCATCTATCCGCTC 2803
Qy 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer 857
Db 2804 TTGGCCTATTGCACATTGCCCTGCCATCGCTGTGTCAGAGGAAATTTATCACGCCAGAG 2863
Qy 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2864 CTTAAACATGTTGCAGCGCTCTGGTTTCATGTCATCTTTCATCTGCATTTTGTCTACGAGC 2923
Qy 878 IleLeuGluMetArgTTPSerGlyValSerIleGluGluTTPTrpArgAsnGluGlnPhe 897
Db 2924 ATCTCGAAATCAGATCGAGTGGTGTAGGCATCGATGATGCTGTGAGAAACGAGCAGTTT 2983
Qy 898 TrpValIleGlyValSerAlaHisLeuPheAlaValGlnGlyLeuLeuVal 917
Db 2984 TGGGTCATGGAGGCGTGTCTTCACATCTCTTGTGTTTCCAGGAGCTCTCTCAAGTTC 3043
Qy 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAsp 937
Db 3044 ATAGCTGTGTAGACACGAGCTTCTATGTGATCCAAAG-----GGCGGAGACGACGAG 3097
Qy 938 GluPheAlaGluLeuTyrAlaPheLysTTPThrThrLeuLeuIleProProThrThrLeu 957
Db 3098 GAGTCTTCAGAGCTGTACACATTCAAATGGACGACCTCTTGATACCTCCGACAAACCTGT 3157
Qy 958 LeuIleIleAsnValIleGlyValAlaGlyIleSerAspAlaIleAsnAsnGlyTyr 977
Db 3158 CTCCTACTGAACCTTCAATGGAGTGGTGGCATCTCAATGGCATCAACACCGATAT 3217
Qy 978 GlnSerTTPGlyProLeuPheGlyLysLeuPhePheAlaPheTTPValIleValHisLeu 997
:::|||||
```

```
Db 3218 GAATCATGGGGCCCCCTGTTCCGGAAGCTCTTCTTTGGCATTTTGGGTGATCGTCCATCTT 3277
Qy 998 TyrProPheIleuysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIle 1017
Db 3278 TACCCCTTCTCAAGGCTCTGGTTGGAGGCGAGAACAGGACGCCAAGATTGTCTATGTC 3337
Qy 1018 TTPSerIleLeuLeuAlaSerIlePheSerLeuLeuTTPValArgIleAspProPheIle 1037
Db 3338 TGGTTCATCTCTCTGCTTCGATCTTCTCGCTGCTTTGGTCCGATCGACCCGTTCTT 3397
Qy 1038 ValArgThrIleGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3398 GCGAAGGATGATGCTCCCTGTTGGAGGAGTGTGGTCTGGATTGC 3442

RESULT 10
US-10-425-114-31130
; Sequence 31130, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCES: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31130
; LENGTH: 3505
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73131F05_FLI
US-10-425-114-31130
```

```
Alignment Scores:
Pred. No.: 0 Length: 3505
Score: 4089.00 Matches: 760
Percent Similarity: 79.55% Conservative: 127
Best Local Similarity: 68.16% Mismatches: 138
Query Match: 72.40% Indels: 90
DB: 17 Gaps: 16
```

US-10-627-132-30 (1-1052) x US-10-425-114-31130 (1-3505)

```
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 84 ATGGAGCCAGCGCGCGGCTGGTCCCGGCTCGCACAAACCGGAACGAGCTGTCGTATC 143
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 144 CGCCGCGATGCGAGCCAGGCGCGAAGCCCATGACAGCGAAGCGGCGAGTGTGCCAG 203
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 204 ATTTGCGCGACGACGCGGGCGCAACCCCGACCGGGAGCGCGTTCGTGGCTTCACACGAG 263
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 264 TGCCTCTTCCCATCTGCGGGGACTGTACAGTACGAGCGCGCGGAGGCGACCGAGAAC 323
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 324 TGCCCCAGTGCAGAACCCCGCTTCAAGCGCTCAAGGGGTGCGCGCGCTGCCCGGGGAC 383
Qy 99 AspAspGluGluAspIleAspAspLeuHisGluPheAsnIleAspAspGluAsnGln 118
Db 384 GAGGAGGAGGAGCGCGCTCGACGACTGTGAGAACGAGATTCAACTGGAGCGACAGCCGAC 443
```

QY 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
DB 444 TCCAGTACCTC-----GCCGAGTCCATGCTCCACGCC 476
QY 139 ArgMetSerTyrGlyArgGlyProAsp---AspGlyAspGlyAsnAsnThrProGlnIle 157
DB 477 CACATGAGTACGGCGCGCGCGACCTCGAGCGGTGCCGAGCCATCCACCCATC 536
QY 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
DB 537 CCC-----AATGTTCCCTC-----CTCACCAAC 560
QY 178 GlyTyrGlyHisGlyGluValSerSerSerLeuHis----- 189
DB 561 GCACAGATGGTCATCACATCCCGCCGACACAGCCCTGTGTGCCCTCGTTCGGGT 620
QY 190 -----LysArgIleHisProTyrProValSerGluProGly-----Ser 202
DB 621 GCGCGGGGAGAGGATTCACCTCTCCCGTAGCGGGATCCCAACCTTCTGTGCACCG 680
QY 203 AlaLysTrpAspGluLysLysGlu-----ValSerTrpLys 214
DB 681 AGGTCTATGGACCTTCCCAAGGATCTCGCCGCATATGGCTACGGGAGCGTAGCATGGAAG 740
QY 215 GluArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----Gly 228
DB 741 GAGAGGATGGAGCTGGAGGAGCAGAGCAGAGGAGATGCACACGAGGAAACGATGCC 800
QY 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
DB 801 GCGCGGATGATGGTAT-----GATGCAGATCTACCACTAATGGATGAAGCTAGACAG 854
QY 249 ProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIle 268
DB 855 CCATTGTCCAGAAAGATCCGCTTCTTCAAGCCAAATCAACCCCTATAGGATGATATA 914
QY 269 ValValArgLeuValValLeuAlaPheLeuArgTyrArgIleLeuHisProValPro 288
DB 915 ATAAATTCGGCTAGTGGTTGTGTTCTTCTTCCACTACCGAGTGATGCGGTGCT 974
QY 289 AspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrp 308
DB 975 GATGATTTGCTTATGCTCATATCTGATCTGTGAATTTGGTTGCCATGCTTGG 1034
QY 309 IleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeu 328
DB 1035 ATCTTGACCACTTCCAAAGTGTTCCTATCGAGAGGAAACCTATCTTCCACCGCTG 1094
QY 329 SerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheVal 348
DB 1095 AGTTTAAAGGTTTGAACAAGAGGGGATCCTTCAACTCGCCCTGTGCGATTTCTTGTG 1154
QY 349 SerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIle 368
DB 1155 AGTACGGTTGATCCCTTGAGGAACCTCATTTGGTCACTGCTAATACTGTTCTATCTATC 1214
QY 369 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspGlyAlaSer 388
DB 1215 CTTTCGGTGGATTATCCAGTTGATTAAGTTTTCATGCTACGTTCTTGATGATGGTCTGCC 1274
QY 389 MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPhe 408
DB 1275 ATGCTGACATTTGAAGCATTTGTGAAACATCTGAATTTGCAAGAAATGGTTCCTTTC 1334
QY 409 CysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAsp 428
DB 1335 TGCAAAAGATATAGCTTGGCTCGCTCCAGAGTGGTACTTCCAAACAGAAATGAGAC 1394
QY 429 TyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGlu 448
DB 1395 TACCTGAAAGACAGGTGGCGCCAAACTTTGTAGAGAACCGAGGAGCAATGAGAGAG 1454
QY 449 TyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAla 468

DB 1455 TATGAGGAATTCAGGTCAAGATCAATGCTTGGTTGCTTAAAGCCCCAAAAGGTTCTCTGAG 1514
QY 469 GluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 488
DB 1515 GAAGATGAGCAATCGAGTGGAACTCCATGCCCGGAAATAATATCCGTGATCATCTCT 1574
QY 489 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeu 508
DB 1575 GGAATGATTCAGGTTTCTTGGTCAAAAGTGTGGCCATGATGTGAAGAAATGAGCTG 1634
QY 509 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAla 528
DB 1635 CCTCATTTGGTTTATGTTTCAAGAGAAAAACGCCGAGCTACACCATCACAAAGAGGT 1694
QY 529 GlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeu 548
DB 1695 GGTGCTATGAATGCATTTGGTCCGAGTCTCTGCTGTACTACTACTACTCTCTTATTTGCTG 1754
QY 549 AsnLeuAspCysAspHisTrpIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhe 568
DB 1755 AACTTTGGATTTGTGATCACATATCAATAATAGTAGGCTATAAAGAGCAATGTGTTTT 1814
QY 569 LeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAsp 588
DB 1815 ATGATGGATTCCTTCTTGGAAATAAAGTTTGTCTATGTGCAGTTTCTCTCAAGATTTGAT 1874
QY 589 GlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPheAspIleAsnMet 608
DB 1875 GGGATGATCGCCATGATCGATATGCTAACAGAAATGTTGCTTTTTCGATATCAACATG 1934
QY 609 LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg 628
DB 1935 AAGGTTTGGATGGTATCCAGGGCCCAATTTATGTGGTACTGGATGTCTTTCAGAGG 1994
QY 629 GlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCys 648
DB 1995 CAGGCATTTATATGGTACGATCTCCCAA---ACAAAGAACCCACCATCAAGAACTTGC 2051
QY 649 Asp-----CysCysProCysPheGly---ArgLysLysArgLys 660
DB 2052 AACTGCTGGCCAAAGTGGTGCATTTGCTGTTGCTGTTTGGTAAACAGGAAAGCAAGAG 2111
QY 661 HisAlaLysAspGlyLeuProGlu----- 668
DB 2112 AAGACCAAGACCTCTAAACCTTAATTTGAGAAGATAAAGAACTTTTAAAGAAAAGGAA 2171
QY 669 -----GlyThrAlaAspMetGlyValAspSer 677
DB 2172 AATCAAGCCCTGCATATGCTTGGTGAATTTGATGAAGCTGCTCCAGGAGCTGAAAT 2231
QY 678 AspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAla 697
DB 2232 GAAAGGCTAGTATTTGTAATCAACAGAAAGTTGGAAGAAATTTGGCCAGTCTTCAGTT 2291
QY 698 PheValThrSerThrLeuMetGluGlyGlyValProProSerSerSerProAlaAla 717
DB 2292 TTTGTTGCATCCACATCTTGAGAAATGGTGAACCTCGAAGAGTGCAGTCCAGCTTCT 2351
QY 718 LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGly 737
DB 2352 CTTCTGAAGGAAGCTATACATGTCTCAGCTGTGATATGAAGACAAACAGGCTGGGA 2411
QY 738 LeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMet 757
DB 2412 AAAGATATTTGGTGGATTTATGGATCAGTCAAGAAAGATATTTTACTGGGTTTAAAGATG 2471
QY 758 HisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777
DB 2472 CACTGCCATGTTGGCGGTCAATTTACTGCTATCTTAACCGGCGCCCTTCAAGGTTCC 2531
QY 778 AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVal 797

Db 2532 GCACCTCTCAATCTTCCGATCGTCTTCCACAGGTTCTTCGGTGGGCTCTTGGTTCAATT 2591
Qy 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrIysAsnGlyAsnLeuLys 817
Db 2592 GAAATTTCTTCAGCAACCACTGCCCTCTCTGGTATGGGTAT---GGTGTGACTAAAG 2648
Qy 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837
Db 2649 TTCTCGAAGAGTTTTCGTACATTAATCCATCGTATACCTTGGACATCTATCCCGCTC 2708
Qy 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyIysPheIleMetProSer 857
Db 2709 TTGGCTTATTGCATATGCTGCCATCTGCTTGTGACAGGAAATTTATCAGCCACAG 2768
Qy 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2769 CTTAAACAATGTGCGACGCTGTGGTTTCATGTCATCTTTCATCTGTCATTTTGTACGAGC 2828
Qy 878 IleLeuGluMetArgTyrSerGlyValSerIleGluGluTyrTrpArgAsnGluGlnPhe 897
Db 2829 ATCTTGAATGAGATGGAGTGTGTAGGCATCGATGCTGTGGAGAAACGACGACTTT 2888
Qy 898 TrpValIleGlyIleValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysVal 917
Db 2889 TGGGTCAATTGGAGCGTGTCTTCATCTCTTTGCTGTGTTCCAGGACTCTCTCAAGGTC 2948
Qy 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerIysAlaThrGlyAspGluAspAsp 937
Db 2949 ATAGCTGGTGTAGACACGAGCTTCACTGTGACATCCCAAG-----GGCGGAGACGACGAG 3002
Qy 938 GluPheAlaGluLeuTyrAlaPheIlystPThrThrLeuLeuIleProProThrThrLeu 957
Db 3003 GAGTTCTCAGAGCTGTACACATTAATGGACGACCTTCTGATACCTCCGACAAACCTTG 3062
Qy 958 LeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyr 977
Db 3063 CTCCTACTGAACCTTCAATTGGAGTGTGTAGCTGGCATCTCCAAATGCGATCAACACGGATAT 3122
Qy 978 GlnSerTrpGlyProLeuPheGlyIysLeuPhePheAlaPheTrpValIleValHisLeu 997
Db 3123 GAATCATGGGGCCCCCTGTTGGGAAGCTCTTCTTGTGATTTGGGTGATCGTCCATCTT 3182
Qy 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIle 1017
Db 3183 TACCGTTCTCTCAAGGTCGTGGTGGGAGGACGACAGACGCCAACGATTTGATGTC 3242
Qy 1018 TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTyrValArgIleAspProPheIle 1037
Db 3243 TGGTCCATCTCTCGCTTCGATCTTCTCGCTGCTTTGGGTCCGATCGACCCGTTCTT 3302
Qy 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3303 GCGAAGGATGATGGTCCCTGTTGGAGGATGTGGTCTGGATTGC 3347

RESULT 11

US-10-425-115-86998
; Sequence 86998, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 86998
; LENGTH: 3783
; TYPE: DNA
; ORGANISM: Zea mays

FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_179348C.1

US-10-425-115-86998

Alignment Scores:
Pred. No.: 0 Length: 3783
Score: 4089.00 Matches: 760
Percent Similarity: 79.55% Conservative: 127
Best Local Similarity: 68.16% Mismatches: 138
Query Match: 72.40% Indels: 90
DB: 18 Gaps: 16

US-10-627-132-30 (1-1052) x US-10-425-115-86998 (1-3783)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 186 ATGAGAGCCAGCCCGGGCTGTGCGGCTCGCACACACCGAAGACGAGCTCGTCGTCATC 245
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 246 CGCGCGATGGCGAGCCAGGGCCGAAGCCCATGCAGCCAGCGAAGCGCCAGGTGTGCCAG 305
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 306 ATTTGGCGCAGCAGCGTGGGGCGCAACCCCGACGGGAGCCGTTCTGGTGGCTGTCAACGAG 365
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 366 TGGCCTTCCCATCTGCGCGGACTGCTACAGTACGAGCGCCGCGAGGCGACGACGAC 425
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 426 TGCCCCCAGTGCAAGACCCGCTTCAAGGGGTGCGCGCGCTGCCCGGGAC 485
Qy 99 AspAspGluLeuAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 486 GAGGAGGAGGAGCGCGCTCGACACCTGAGAGAGTTCACTGAGCGGACGACGACGAC 545
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 546 TCCAGTACCTC-----GCCGAGTCCATGCTCCACGCC 578
Qy 139 ArgMetSerTyrGlyArgGlyProAsp---AspGlyAspGlyAsnAsnThrProGlnIle 157
Db 579 CACATGAGTACGCGCGCGCGCGACCTCGACGCGCTGCGCGACCATTCCTCACCCCATC 638
Qy 158 ProProIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
Db 639 CCC-----AATGTTCCCTC-----CTCACCAAC 662
Qy 178 GlyTyrGlyHisGlyGluValSerSerSerLeuHis----- 189
Db 663 GGACAGATGGTGCATGACATCCCGCCGAGCAGCACGCCCTTGTGCTCGTGGGT 722
Qy 190 -----LysArgIleHisProTyrProValSerGluProGly-----Ser 202
Db 723 GGCGGGGGAAGAGATTACCTCTCCGTCAGCGGATCCCACTTCCTCTGTGCAACCG 782
Qy 203 AlaLysTrpAspGluLysLysGlu-----valSerTrpLys 214
Db 783 AGGTCTATGGACCTTCCAAAGGATCTCGCCCATATGGCTACGGGAGCGTAGCTGAAG 842
Qy 215 GluArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----Gly 228
Db 843 GAGAGATGGAGAGCTGGACGACAGCAGGAGGAGGATGACACGAGGAGGACGATGCG 902
Qy 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
Db 903 GGCGCGCATGATGTTGAT-----GATGCAGATCTACCACTTAATGATGAAGTAGACAG 956
Qy 249 ProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIle 268
Db 957 CCATTGTCCAGAAAGATCCCGCTTCTTCAAGCCAAATCAACCCCTTAGATGATTATA 1016


```

; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-49

Alignment Scores:
Pred. No.: 0 Length: 3969
Score: 4073.50 Matches: 757
Percent Similarity: 79.68% Conservative: 129
Best Local Similarity: 68.08% Mismatches: 139
Query Match: 72.12% Indels: 87
DB: 16 Gaps: 15

US-10-627-132-30 (1-1052) x US-10-209-059-49 (1-3969)

Qy 1 MetGluAaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu1e 20
Db 144 ATGAGCGCAGCGCCGGCGTGGTGGCGGCTCCACACCGCAAGAGTGTGTGTCATC 203
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 204 CGCCGCGACGGCCATCCCGGCGCGAGCGCGCGGCGGACGACGAGCGGCGTGGCCAG 263
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 264 ATTTCGGCGACGAGTGGCGCTTGGCCCGCGCGGCGGACCCCTTCTGTGGCGTGCACGAG 323
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGlyThrGlnAsn 78
Db 324 TGGCGCTTCCCGTCTGCGCGGACTGCTACGAATACGAGCGCGGCGGCGACGAGAAC 383
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 384 TGCCCCCAGTGCAGACTCGATCAAGCGCTCAAGGGCTGCCACGCTGTGACCGGTGAC 443
Qy 99 AspAspGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 444 GAGGAGGAGGAGCGCGTGCATGACCTGGACACAGAGTTCAACTGGGACGCCATGACTCG 503
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 504 CAG-----TCTGTGGCGGAGTCCATGCTCTACGGC 533
Qy 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnThrProGlnIlePro 158
Db 534 CACATGAGCTACGGCCGTGA-----GGTGACCCTAATGGCGCGCCACAGCTTTC 584
Qy 159 ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
Db 585 CAGCTC-----AACCCTAATGTTCCACTC-----CTCACCAACGGG 620
Qy 179 TyrGlyHisGlyGluValSerSerLeuHis----- 189
Db 621 CAAATGTGTGGATGATACATCCACCGGAGCAGCAGCGCTGTGTGCTTCTTTTCATGGGTGT 680
Qy 190 -----LysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db 681 GGGGGAAGAGGATACATCCCTTCTTATTCGGATCCAGCTTACCTGTGCAACCCAGG 740
Qy 204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Db 741 TCTATGGACCCATCCAAGGATCTGTGTCATATGGGTATGGTAGTGTGCTTGGAGGAA 800
Qy 216 ArgMetAspAspTrpLysSerLysGln-----GlyIleLeuGlyGly 229
Db 801 CGGATGGAGAAATTGGAACGACAGACAAGAGAGGATGCACACGCGGGAATGATGTGTGT 860

```


Qy	39	IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValalaCysAsnGlu	58
Db	264	ATTTCGGCGACGACGTCGAGCTTCGCCCTTCGCCGGCGGGACCCCTTCGTGGCGTGCAACGAG	323
Qy	59	CysGlyPheProValCysAspArgProCysTyrGlyTyrGluArgArgGlyGlyThrGlnAsn	78
Db	324	TGGCGCTTCGCCGCTCTCGCGGACCTGCTACGAATACGAGCGCGGAGGCGCACGAGAAC	383
Qy	79	CysProGlnCysIysThrArgTyrLysAsrGluLeuLysGlySerProArgValalaGlyAsp	98
Db	384	TGCCCCAGTGCAGACTCGATCAAGCGCTCAAGGCTGCGACAGTGTGACCGGTGAC	443
Qy	99	AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln	118
Db	444	GAGGAGGAGCAGCGCGTCGATGACCTGGACAACGAGTTCACACTGGGACGCGCCACTGCG	503
Qy	119	GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly	138
Db	504	CAG-----TCTGTGGCCGAGTCCATGCTCTACGGC	533
Qy	139	ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro	158
Db	534	CACATGAGCTACGCGCGTGA-----GGTGACCTAATGGCGCGCACAGCTTTC	584
Qy	159	ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly	178
Db	585	CAGTCTC-----AACCCCAATGTTCCACTC-----CTCACCAACGCG	620
Qy	179	TyrGlyHisGlyGluValSerSerSerLeuHis-----	189
Db	621	CAAAATGCTGGATGACATCCACCGGACGACGCGCTGGTGCTCTCTTCATGGGTGT	680
Qy	190	-----LysArgIleHisProTyrProValSerGluProGly-----SerAla	203
Db	681	GGGCGGAAGAGGATACATCCCTCTTATGCGGATCCGAGTACTGCTGCAACCCAGG	740
Qy	204	LysTrpAspGluLysLysGlu-----ValSerTrpLysGlu	215
Db	741	TCATATGACCCATCCAGGATCTTGCTGCATATGGGTATGGTATGTTGCTTGGAAAGAA	800
Qy	216	ArgMetAspAspTrpLysSerLysGln-----GlyIleLeuGlyGly	229
Db	801	CGGATGAGAATTCGAAGCAGACAGACAGAGAGGATGCACCAGACGGGGAATGATGTTGT	860
Qy	230	GlyAlaAspProGluAspMetAspAlaAspValalaLeuAsnAspGluAlaArgGlnPro	249
Db	861	GGTGATGATGGTGAC-----GATGCTGATCTACCACTAATGGATGAAGCAACACAA	914
Qy	250	LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal	269
Db	915	CTGTCCAGAAAATTCACCTTCATCAAGCCAGATTAATCCATATAGATGATATCAATT	974
Qy	270	ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp	289
Db	975	ATTTCGGCTTGTGGTTTGGGGTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGAT	1034
Qy	290	AlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIle	309
Db	1035	GCATTTGCTTTGTGGCTCATATCTGTAATCTGTGAAATCTGGTTGGCATCTGTGGANT	1094
Qy	310	LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSer	329
Db	1095	CTTGATCAATTCCTCCNAAGTGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGCA	1154
Qy	330	LeuArgTyrGluArgGlyGlyProSerLeuLeuSerAlaValAspLeuPheValSer	349
Db	1155	CTGAGTTTCGACAGGAAGGCCAGCCATCTCAACTTGCTCCAAATGATTTCTTTGTCACT	1214
Qy	350	ThrValAspProLeuLysGluProLeuValThrAlaAsnThrValLeuSerIleLeu	369
Db	1215	ACGGTTGATCCCTTAAGAGAACCTCTCTTTGGTGCAACAATACTGTTCTATCTATCTCT	1274
Qy	370	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	389

[illegible]

Db 2352 TCCACACTTCTCGAGAAATGGTGGACCTTGAAGAGTGCAGTCCCTCTCTCTTTTGAAA 2411
 Qy 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
 Db 2412 GAAGCTATATACATGTCATTAAGTGTGTTATGAAGACAGACAGACTGGGCAAGAGATT 2471
 Qy 741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
 Db 2472 GGCTGATCTATGGATCACTTACAGAGATATCTTAACCTGGTTCAAGATGATGTCAT 2531
 Qy 761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
 Db 2532 GGTGGCGGTCAATTACTTGCATACCTAAACGGGTGTCATTCAAGGTTCTGCACCTCG 2591
 Qy 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
 Db 2592 AATCTTTTCAGATCGCTTCCACAGGTGCTTCGGTGGCTCTTGGGTCTATTGAGATCTTC 2651
 Qy 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
 Db 2652 TTCAGCAATCATGGCTCTTGGTATGGGTAT---GGTGGCGGTCTGAATAATTTTGGAA 2708
 Qy 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
 Db 2709 AGATTTCTCTACATCACTCCATCGTATCTCTGGACATCTATCCCTCTTGGCTTAC 2768
 Qy 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
 Db 2769 TGTATACATGGCTGCCTCTTGTATTTATGACAGGAAATTTATCACTCCAGAGCTGAATAAT 2828
 Qy 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
 Db 2829 GTTGCAGCTGTGGTGTATGTCATCTTTATCTGCACTTTTATCTGCACTTTTGTACGAGCATCTAGAA 2888
 Qy 881 MetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluGlnPheTrpValIle 900
 Db 2889 ATGAGATGGAGTGGTGGTGGATTTGATGCTGTGGAGGAAATGACAGATTCTGGGTCAIT 2948
 Qy 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLysValLeuAlaGly 920
 Db 2949 GGAGGTGTCTCTCACACCTCTTGTGTGTTCCAGGACTTCTCAAGTCTATGCTGT 3008
 Qy 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAspGluPheAla 940
 Db 3009 GTTGATACAAAGCTTCAACCTGATCAAAAG-----GGTGGAGATGATGAGGAGTTCTCA 3062
 Qy 941 GluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProThrThrLeuLeuIleIle 960
 Db 3063 GAGCTATATACATTCAATGGACTACCTTATTGATACCTCTCTACACCTTGTCTTATATG 3122
 Qy 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
 Db 3123 AACTTCATTGGTGGTGGCTGGCTTCAAAATGCGATCAATAACGATATGAGTCATGG 3182
 Qy 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1000
 Db 3183 GGCCCTCTCTTGGGAAGCTATCTTGGCATTTGGGTGATGTCATCTTTATCCCTTT 3242
 Qy 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
 Db 3243 CTCAAAGGTTGGTGGAGGCAAAACAGACCAACAGATGTCATGCTGTGGTCAAT 3302
 Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
 Db 3303 CTGCTGGCTTCAATCTCTCGCTCTTGGGTGGGATGATGATCTTCTTCTTCCGAAGAT 3362
 Qy 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db 3363 GATGTCCTGCTCTTGGAGAGTGTGGTGGATTGC 3398

RESULT 15

US-10-160-719-37

; Sequence 37, Application US/10160719

; Publication No. US20030167528A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kanwarpal S. Dhugga
 ; APPLICANT: Timothy G. Helentjaris
 ; APPLICANT: Benjamin A. Bowen
 ; APPLICANT: Xun Wang
 ; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
 ; FILE REFERENCE: Thereof
 ; CURRENT APPLICATION NUMBER: US/10/160,719
 ; PRIOR FILING DATE: 2002-06-03
 ; PRIOR APPLICATION NUMBER: US/09/371,383A
 ; PRIOR FILING DATE: 1999-08-06
 ; PRIOR APPLICATION NUMBER: US 60/096,822
 ; PRIOR FILING DATE: 1998-08-17
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 37
 ; LENGTH: 3969
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (144)....(3401)
 US-10-160-719-37

Alignment Scores:
 Pred. No.: 0 Length: 3969
 Score: 4073.50 Matches: 757
 Percent Similarity: 79.68% Conservative: 129
 Best Local Similarity: 68.08% Mismatches: 139
 Query Match: 72.12% Indels: 87
 DB: 16 Gaps: 15

US-10-627-132-30 (1-1052) x US-10-160-719-37 (1-3969)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
 Db 144 ATGGAGGCGAGCGCGGCTGTGGCTCCACAAACCGCAACGAGCTGCTGCTCATC 203
 Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
 Db 204 CGCCGCGACGCGCATCCCGGCGGAGCCGCGGGAGAGAGACGCGGAGCGGTGTGCCAG 263
 Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
 Db 264 ATTTGGCGGACGACGTCGGCTTCCCGCGGGGGACCCCTTCTGTCGCTGCAACGAG 323
 Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
 Db 324 TGCCTTCCCGCTCTGCGGAGCTGTCTACGAATACGAGCGCGGAGGCGACGAGAAC 383
 Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
 Db 384 TCCCCCAGTCCAAAGCTCGATACAGCGCTCAAGGCTCCCAACGTCGTCACCGGTGAC 443
 Qy 99 AspAspGluGluAspIleAspAspLeuHisGluPheAsnIleAspAspGluAsnGln 118
 Db 444 GAGGAGGAGGACGCGTCGATGACTGGACAAACAGATTCACTGGGACGCGCATGACTCG 503
 Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
 Db 504 CAG-----TCTGTGGCGGAGTCCATGCTCTACGCG 533
 Qy 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAsnAsnThrProGlnIlePro 158
 Db 534 CACATGAGTACGCGCGGTGA-----GTCACCTAATGGCGGCCCAAGCTTTC 584
 Qy 159 ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
 Db 585 CAGCTC-----AACCCCAATGTTCCACTC-----CTCACCACACGGG 620
 Qy 179 TyrGlyHisGlyGluValSerSerLeuHis----- 189

Db 621 CAAATGGTGGATGACATCCACCGGAGCAGCAGCGCTGTGGCTTCTTTCATGGGTGGT 680
Qy 190 -----LysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db 681 GGGGAAAGAGGATACATCCCTTCTTATGCGGATCCACGCTTACCTGTGCAACCCAGG 740
Qy 204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Db 741 TCTATGACCCATCAAGGATCTTGTGTCATATGGGTATGGTAGTGTGTCTTGGNAGGAA 800
Qy 216 ArgMetAspTrpLysSerLysGln-----GlyIleLeuGlyGly 229
Db 801 CGGATGAGAAATTCGAGCAGACAGAGAGGATGCACACGCGGGAATGATGGTGGT 860
Qy 230 GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaAArgGlnPro 249
Db 861 GGTGATGATGGTGAC-----GATGCTGATCTACCACTAAATGGATGAAGCAACAA 914
Qy 250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
Db 915 CTGTCCAGGAAATTCACCTTCCATCAAGCCAGATTAATCCATATAGGATGATATCAT 974
Qy 270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
Db 975 ATTCGGCTTGTGGTCTTCTTCTTCCACTACCGAGTGATGCCGTGAATGAT 1034
Qy 290 AlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIle 309
Db 1035 GCATTTGCTTTGTGGCTCATATCTGTATCTGTGAATCTGGATCTGTGGAT 1094
Qy 310 LeuAspGlnPheProLysTrpPheProIleAspArgGluTrpTyrLeuAspArgLeuSer 329
Db 1095 CTTGATCAATTCCTCAAGTGGTTCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCA 1154
Qy 330 LeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
Db 1155 CTGAGTTTCGACAGGAAGCCAGCCATCTCAACTTGTCTCAATTTGATTTCTTGTGAGT 1214
Qy 350 ThrValAspProLeuLysGluProProLeuValThrAlaAnThrValLeuSerIleLeu 369
Db 1215 ACGGTGTGATCCCTTAAAGGAACTCTTGTGTCACAAAATACCTGTTCTATCTATCTCT 1274
Qy 370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 389
Db 1275 TCGGTGGATATTCCTGTTGATAAGGTTTCTTGTCTANGTTCTGATGATGGTGTGCAATG 1334
Qy 390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys 409
Db 1335 CTAAGCTTTGAAGCATTATCTGAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGTC 1394
Qy 410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
Db 1395 AAACGGTACAATTAATTTGAACCTCGCGCTCCAGAGTGGTACTTTCCAAACAGAGATAGACTAC 1454
Qy 430 LeuLysAspLysValGlnProThrPheValGlnArgAtqAlaMetLysArgGluTyr 449
Db 1455 TTGAAGAACAGGTGGCGCAACCTTTGTTAGGAGAGAGAGCAATGAGAGAGAGAT 1514
Qy 450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
Db 1515 GAGGAATTCAGGTGAGAATCAATGCCTTAGTTGCCAAGGCCAGAAAGTTCTCTGAAGAA 1574
Qy 470 GlyTrpIleMetLysAspGlyThrProTrpProGlyYasnAnThrArgAspHisProGly 489
Db 1575 GGATGGACAAATCAAGATGGAAACCCCTGGCCCTGGAAACAAATGTTCTGATCATCTCGA 1634
Qy 490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
Db 1635 ATGATTCAGGCTCTTCTTGGCCAAAGCGAGGCGCTTGACTGTGAGGGGAAATGAATGCA 1694
Qy 510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 529

Db 1695 CGATTCGTTATGTTTCTAGAGAGAAACGACCGAGGCTATAACCATCATTAAGAAAGCTGGT 1754
Qy 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
Db 1755 GCTATGATGATGCTGTGCGAGTCTCTGCTGTACTAAACAAATGCTCCATATTTGTTAAAC 1814
Qy 550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
Db 1815 TTGGATTTGTGATCACTACATCAACAACAGCAGCTATAAAGGAAGCAATGTGTTTATG 1874
Qy 570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
Db 1875 ATGGACCTTTTACTAGGAAGAAGGTTTGTATGTACAGTTCCTCAAGATTTGATGGG 1934
Qy 590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
Db 1935 ATTGATCGCCATGACCGATATGTAACCGAATGTGTCTTTTGTATATCAACATGAAA 1994
Qy 610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln 629
Db 1995 GGTTTGGATGGTATTCAGGGTCCAAATTTATGTTGGTACTGATGTGTATTAGAAAGCAG 2054
Qy 630 AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp 649
Db 2055 GCATTTATGTTTATGATGCCCTCCCAAA---ACAAGAAGCCCATCAAGGACTTGCAC 2111
Qy 650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
Db 2112 TGCTGGCCCAAGTGGTCTTGTCTGTCTGCTTTGGCAATAGGAAGCAAAAGAGACT 2171
Qy 663 LysAspGlyLeuProGlu----- 668
Db 2172 ACCAAACCCAAACAGAGAGAAAGTATTTATTTTCAAGAAAGAGAACCAATCC 2231
Qy 669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
Db 2232 CCTGCAATATCTCTTGTGTAATTCAGAAAGCTCTCCAGAGCTGAGATGAAGAAGCC 2291
Qy 681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
Db 2292 GGTATGTTAATCNACAAAATTTAGAAAAGAAATTTGGCCAATCTCTCTTTTGTATACA 2351
Qy 701 SerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuLys 720
Db 2352 TCCACACTTCTCGAGAATGGTGAACCTTGAAGAGTGCAAGTCTCTCTCTTTTGA 2411
Qy 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
Db 2412 GAAGCTATACATGTCATTAGTTGGTTATGAAGACAGACAGACTGGGGAAGAGAGAT 2471
Qy 741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
Db 2472 GGCTGGATCTATGATCAGTTACAGAGATATTTCTAATCTGGTTCAAGATGCTATGTCAT 2531
Qy 761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
Db 2532 GGTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTG 2591
Qy 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
Db 2592 AATCTTTTCAGATCTCTCTCCAGGCTCTCGGTGGGCTCTTGGGTCTTATTGAGATCTTC 2651
Qy 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyYasnLeuLysTrpLeuGlu 820
Db 2652 TTGAGCAATCATGCCCTCTTGGTATGGGTAT---GGTGGCGGTCTGAATTTTGGAA 2708
Qy 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
Db 2709 AGATTTTCTACATCAACTCCATCGTGATCTTGGACATCTATTCCCTCTTGGCTTAC 2768
Qy 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTACATTCCTGCCATCTGTTTATTGACAGGGAATTTATCATCCAGAGCTGATTAAT 2828


```
QY      861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db      2829 GTTGCAGCCTGTGGTTCATGTCACATTTTATCTGCATTTTGTACGAGCATCCTAGAA 2888
QY      881 MetArgTyrSerGlyValSerIleGluGluTyrTrpArgAsnGluGlnPheTrpValIle 900
Db      2889 ATGAGATGGAGTGGTGTGGATTTGATGACTGGTGGAGAAATGAGCAGTTCTGGGTCAAT 2948
QY      901 GlyGlyValSerAlaHisLeuPheAlaValAlaGlnGlyLeuLeuLysValLeuAlaGly 920
Db      2949 GGAGGTGTGCTCCACACCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCAATAGCTGCT 3008
QY      921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
Db      3009 GTTGATACAAAGCTTCCCGTGACATCAAAAG-----GGTGGAGATGATGAGGAGTTCTCA 3062
QY      941 GluLeuTyrAlaPheLysTyrThrThrLeuLeuIleProThrThrLeuLeuIleIle 960
Db      3063 GAGCTATATACATTCAAATGGACTACCTTATTGATACCTCTACCACCTTGCTTCTATTG 3122
QY      961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
Db      3123 AACTTCATTGGTGTGGTCTGGCTTCAATGCGATCAATAACGGATATGAGTCATGG 3182
QY      981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1000
Db      3183 GGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTGGGTGATTGTCCATCTTTATCCCTTT 3242
QY      1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
Db      3243 CTCAAAGGTTTGGTGGAGGCAAAACAGCACCAACGATTTGTCATGCTGTGGTCCATT 3302
QY      1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db      3303 CTGCTGGCTTCAATCTTCTCGCTCTTTGGGTTCGATTGATCTTTCTTCGGAAGGAT 3362
QY      1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db      3363 GATGCTCCGCTTCTTGGAGAGTGTGGTTTGGATTGC 3398
```

Search completed: March 12, 2005, 16:29:51

Job time : 1045 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2005, 00:02:06 ; Search time 9650 Seconds
(without alignments)
17288.218 Million cell updates/sec

Title: US-10-627-132-29
Perfect score: 3443
Sequence: 1 ctgcgtgcctgcctgcga.....aaaaaaaaaaaaaaaaaaaaa 3443

Scoring table:  Gapop 60%0.5, Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues



Total number of hits satisfying chosen parameters: 226460

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3443	100.0	3443	8	AY372246	AY372246 Zea mays
2	83	2.4	3631	8	AK121170	AK121170 Oryza sat
3	79	2.3	165909	8	AF005420	AF005420 Oryza sat
4	79	2.3	187410	8	AF005579	AF005579 Oryza sat
5	59	1.7	1255	8	AY483156	AY483156 Hordeum v
6	50	1.5	3470	8	AY372244	AY372244 Zea mays
7	49	1.4	814	5	BC058699	BC058699 Danio rer
8	49	1.4	4919	9	HSM080426	HSM080426 Homo sapi
9	48	1.4	1368	10	BC03654	BC03654 Rattus no
10	48	1.4	1792	9	BC017482	BC017482 Homo sapi
11	48	1.4	13981	9	AC141002	AC141002 Homo sapi
12	47	1.4	431	6	CQ700614	CQ700614 Sequence
13	47	1.4	1409	10	BC058737	BC058737 Mus muscu
14	47	1.4	1417	5	ZOPY18678	ZOPY18678 Zosterisess
15	47	1.4	1431	9	BC039586	BC039586 Homo sapi
16	47	1.4	1450	10	BC019205	BC019205 Mus muscu
17	47	1.4	1897	10	BC048832	BC048832 Mus muscu
18	47	1.4	2094	9	BC043552	BC043552 Homo sapi
19	47	1.4	2513	10	BC018554	BC018554 Homo sapi

c 93 1.3 132296 10 AL772283 AL772283 Mouse DNA
c 94 1.3 149567 10 AC136639 AC136639 Mus muscu
95 1.3 149994 10 AC147259 AC147259 Mus muscu
96 1.3 152944 10 AC144580 AC144580 Mus muscu
c 97 1.3 154136 2 AC113980 AC113980 Mus muscu
c 98 1.3 158469 10 AC145586 AC145586 Mus muscu
99 1.3 165604 9 AC008927 AC008927 Homo sapi
c 100 1.3 188444 10 AC150035 AC150035 Mus muscu

ALIGNMENTS

RESULT 1
AY372246 3443 bp mRNA linear PLN 01-SEP-2004
LOCUS Zea mays cellulose synthase catalytic subunit 12 (CesA12) mRNA,
DEFINITION complete cds.
ACCESSION AY372246
VERSION AY372246.1 GI:38532103
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3443)
AUTHORS Dhugga,K.S., Barreiro,R., Appenzeller,L., Wang,H., Niu,X.,
Carriagan,L. and Tomes,D.

Cellulose formation and its role in determining stalk strength in
maize

Unpublished
2 (bases 1 to 3443)
AUTHORS Dhugga,K.S.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-2003) Trait and Technology, Pioneer Hi-Bred
International, Inc., a Dupont company, 7300 NW 62nd Avenue,
Johnston, IA 50131, USA

FEATURES
Location/Qualifiers
1..3443
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
1..3443
/gene="CesA12"
/note="CESA12; ZmCESA12"

gene

CDS

/codon_start=1
/product="cellulose synthase catalytic subunit 12"
/protein_id="AA23312.1"
/db_xref="GI:38532104"
/translations="MEASAGLVASHNRNELVIRGHEDPKLRALSQVCEICGDEV
GLTVGDLFVACNECPVPCVYERREGTNCQCKTRYKRLKGSFRVAGDDDEE
DIDDLHEFNIDENQRLGEMQNSQITEAMLRHMSYRSGPDGDGNTQPIPP1
ITGSRVPSVSGEPFITTPYGHGEVSSSLHKRIHPYPSBPSAKMDKEKESWKERM
DKSKGILGGADPDMDALNDEARQPLSRKVSIASSKVPYRMVIVVRLVLA
FFLYRLHPVDAIGLTVSIICEIWFALISWLDQFPKWFPIIDRETYLDRLSRYR
EGEPLSLAFAVRLVSTVDPLKEPPLVANTVLSILADVDPVKVSCVSDDGASMLTF
ESLSEAFAPKRVKPKKFGIEPRAPFYSLKVDYLDKDKVQPTFVQERRAMKREYE
EFKVRNALVAKAMKPAEGIMKQTPMPNTRDHPGMIQVFLGHSGHDTGEGNEL
FLVYVYREKRGFQHHKAGANMALIRVSAVLTNAPFMLNDCDHYINNSKAIREAM
CFMLDQVQRKVCYQVFPORFDGIDVDRYANRNTVFDDINMLKLDGIQGPVYVGTG
VFRQALYGNRPKRPKPMVTDCCPCFGKRRKHAKDGLPEGVADMGVSDKEML
MSHMFKEKRGQSAAFVSTLMEEGVVPSSPAALLKEAIIHVISCGYEDKTDWGLEL
GWYIGSITEDLLTGFKMHCGRWSVYCMKRAAPKGSAPINLSLDRLANQVLRWALGSVE
IFFSRSLPFIYGNKLNKLRFAINTIITYPTSLPLAYCTLPVACLLTKGFLMP
SISTFASLFIALFNFISFATGFILEMWRSVGSIEBWRNEQFWVIGVSAHLFAVQGL
LKVLAGIDNFTVTSKATGDEDDFAELYAFKWTLLIIPPTILLIINVIQVWAGISDA
INNGYQWGPFLFKLPFAFWIVLHLYPFLKGLMGRNRTPTVVVWISILLASIFSLW
VRIDPFIVRTKGPDRVQCINGC"

ORIGIN

Query Match 100.0%; Score 3443; DB 8; Length 3443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCGCTCGCCCTCGCTCGCAATCGCGAAATCTGTCGAGCACCTGAGGGTTCGAGGCGCGAG 60
DB 1 CTGCGTGCCTCTCGCTCGCAATCGCGAAATCTGTCGAGCACCTGAGGGTTCGAGGCGCGAG 60
QY 61 AGCTAGCTTAGCACGCGCGCTCTCGCGCGCATGAGGCGCAGCGCGGGCTGCTGCTGCCCGG 120
DB 61 AGCTAGCTTAGCACGCGCGCTCTCGCGCGCATGAGGCGCAGCGCGGGCTGCTGCTGCCCGG 120
QY 121 CTCGCAAAACCGGAAACGAGCTGTGTGATTCGCGGGCCACGAGGACCCCAAGCCGCTGCG 180
DB 121 CTCGCAAAACCGGAAACGAGCTGTGTGATTCGCGGGCCACGAGGACCCCAAGCCGCTGCG 180
QY 181 GGGCTGAGCGGGGAGGTGTGAGATATGCGGCGAGAGGTGCGGCTCACGGTGACGCG 240
DB 181 GGGCTGAGCGGGGAGGTGTGAGATATGCGGCGAGAGGTGCGGCTCACGGTGACGCG 240
QY 241 CGACCTCTTCTGCTGCGCTGCAACGAGTGCCTTCCCGGTGTGCCGCCCTGCTACGAGTA 300
DB 241 CGACCTCTTCTGCTGCGCTGCAACGAGTGCCTTCCCGGTGTGCCGCCCTGCTACGAGTA 300
QY 301 CGAGCGCCGGGAGGCGACGCAAACTGCCCCAGTGCAAGCGCTACAAGCGCTCAA 360
DB 301 CGAGCGCCGGGAGGCGACGCAAACTGCCCCAGTGCAAGCGCTACAAGCGCTCAA 360
QY 361 GGGGAGCCCGAGGTTGCGGGGACGATGACGAGGAGCATGACGAGCTGGAGCACGA 420
DB 361 GGGGAGCCCGAGGTTGCGGGGACGATGACGAGGAGCATGACGAGCTGGAGCACGA 420
QY 421 GTTCAACATCGACGACGAGATCAGCAGAGGAGCTGGAGGGCAACATGACAGAACAGCCA 480
DB 421 GTTCAACATCGACGACGAGATCAGCAGAGGAGCTGGAGGGCAACATGACAGAACAGCCA 480
QY 481 GATCACGAGGCGATGCTGCAAGGAGATGAGCTACGAGGAGGGGCGCGACGAGCGGA 540
DB 481 GATCACGAGGCGATGCTGCAAGGAGATGAGCTACGAGGAGGGGCGCGACGAGCGGA 540
QY 541 CGGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGGCTGAG 600
DB 541 CGGCAACACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGGCTGAG 600
QY 601 CGGTGAGTTTCCGATTACCAACGGGTATGCGCCACGCGGAGGTCTCGTCTTCTCCCTGCACAA 660
DB 601 CGGTGAGTTTCCGATTACCAACGGGTATGCGCCACGCGGAGGTCTCGTCTTCTCCCTGCACAA 660
QY 661 GCGCATCCATCCGTACCTGTGTGAGCCAGGAGTCCAAAGTGGGACGAGAGAAAGA 720
DB 661 GCGCATCCATCCGTACCTGTGTGAGCCAGGAGTCCAAAGTGGGACGAGAGAAAGA 720
QY 721 AGTGAGCTGGAAGGAGGATGAGACGACCTGGAAGTCCAAAGCAGGCGATCCTCGCGCGCG 780
DB 721 AGTGAGCTGGAAGGAGGATGAGACGACCTGGAAGTCCAAAGCAGGCGATCCTCGCGCGCG 780
QY 781 CGCCGATCCCGAAGAATGGAAGCGCCGACCTGGCACTGAACGAGCGGAGGCGAGCGCT 840
DB 781 CGCCGATCCCGAAGAATGGAAGCGCCGACCTGGCACTGAACGAGCGGAGGCGAGCGCT 840
QY 841 GTGAGAGAAAGGTGTCGATCGCGTTCGAGCAAGGTGAACCCGCTACCGATGGTGTGCTGGT 900
DB 841 GTGAGAGAAAGGTGTCGATCGCGTTCGAGCAAGGTGAACCCGCTACCGATGGTGTGCTGGT 900
QY 901 GGGTCTCGTGTGCTCGCTCTTCTCCGCTACCGTATCTGCAACCCGCTCCCGGACGC 960
DB 901 GGGTCTCGTGTGCTCGCTCTTCTCCGCTACCGTATCTGCAACCCGCTCCCGGACGC 960
QY 961 CATCGGCTGTGCTGCTCTCCATCATCTGCGAGATCTGTTGCGCATCTCTCTGGATCCT 1020
DB 961 CATCGGCTGTGCTGCTCTCCATCATCTGCGAGATCTGTTGCGCATCTCTCTGGATCCT 1020
QY 1021 CGACCAAGTTCGCCCAAGTGGTTCGCCCATGACCGCGAGAGCTACCTCGACGCGCTCTCCCT 1080

[illegible]

Db	2101	CGCTGATATGGGAGTAGATAGCGCAACGAGAGATGCTCATGTCCCAATGAACATTCAGAA	2161
Qy	2161	CGCGTTTCGGGCAGTCCGCGGTTTCGTTACAGTCAGCGCTCATGAGAGAAAGCGCGTCCC	2220
Db	2161	CGCGTTTCGGGCAGTCCGCGGTTTCGTTACAGTCAGCGCTCATGAGAGAAAGCGCGTCCC	2220
Qy	2221	TCCTTCGTCGAGCCCCCGCGCTCTCTCAAAGGAGGCATCCATGTTCATCAGCTGCGCCTA	2280
Db	2221	TCCTTCGTCGAGCCCCCGCGCTCTCTCAAAGGAGGCATCCATGTTCATCAGCTGCGCCTA	2280
Qy	2281	CGAGGACAAACCGACTGGGGGCTGGAGCTGGGGTGGATCTACGGGTGCATACGAGGA	2340
Db	2281	CGAGGACAAACCGACTGGGGGCTGGAGCTGGGGTGGATCTACGGGTGCATACGAGGA	2340
Qy	2341	CATCCTGACGGGGTTCAAGATGCACTGCCGCGGTGGCGTCCGTGTACTGCAATGCCGAA	2400
Db	2341	CATCCTGACGGGGTTCAAGATGCACTGCCGCGGTGGCGTCCGTGTACTGCAATGCCGAA	2400
Qy	2401	CGCGGCGCGGTTCAAAGGGTTCGGCGCGATCAATCTATCGGACCGTCTCAACAGGTGCT	2460
Db	2401	CGCGGCGCGGTTCAAAGGGTTCGGCGCGATCAATCTATCGGACCGTCTCAACAGGTGCT	2460
Qy	2461	CCGTTGGGCGTGGGGTTCGTCGAGATCTTCTTCAGCGGCAACGCCCCCTGCTGTAACG	2520
Db	2461	CCGTTGGGCGTGGGGTTCGTCGAGATCTTCTTCAGCGGCAACGCCCCCTGCTGTAACG	2520
Qy	2521	CTACAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCTACATCAACACCAACCATCTA	2580
Db	2521	CTACAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCTACATCAACACCAACCATCTA	2580
Qy	2581	CCCCCTTCACTTCGCTCCGCTGTCTGCGCTACTGCACCCCTCCCGCGCGTCTGCCTCTC	2640
Db	2581	CCCCCTTCACTTCGCTCCGCTGTCTGCGCTACTGCACCCCTCCCGCGCGTCTGCCTCTC	2640
Qy	2641	CGGCAAGTTTCATCATGCGCTGCATTCAGCACTTCGCGAGCCTCTTCTCATCGCCCTCTT	2700
Db	2641	CGGCAAGTTTCATCATGCGCTGCATTCAGCACTTCGCGAGCCTCTTCTCATCGCCCTCTT	2700
Qy	2701	CATGTCATCTTCGCGACGGGCATCTTCGAGATGCGGTGGAGCGGTGAGCATCAGGA	2760
Db	2701	CATGTCATCTTCGCGACGGGCATCTTCGAGATGCGGTGGAGCGGTGAGCATCAGGA	2760
Qy	2761	GTGTTGAGGAACGAGCAGTCTCTGGGTTCATTCGCGCGGTGTCGCGCATCTCTTCGCCGT	2820
Db	2761	GTGTTGAGGAACGAGCAGTCTCTGGGTTCATTCGCGCGGTGTCGCGCATCTCTTCGCCGT	2820
Qy	2821	CGTCAGGGCTTCGTTCAAGTCTTCGCGGGATCGACCAACTTCACGTCACTCTCCAA	2880
Db	2821	CGTCAGGGCTTCGTTCAAGTCTTCGCGGGATCGACCAACTTCACGTCACTCTCCAA	2880
Qy	2881	GGCCACCGGCGAGGAGGACGACGATTCGCGGAGCTACGCCCTTCAAGTGGACCAAGCT	2940
Db	2881	GGCCACCGGCGAGGAGGACGACGATTCGCGGAGCTACGCCCTTCAAGTGGACCAAGCT	2940
Qy	2941	CCTCATCCCCCGCCACCACTGCTCATCTAATTAACGTTCATCGGCGTTCGCGCATCTC	3000
Db	2941	CCTCATCCCCCGCCACCACTGCTCATCTAATTAACGTTCATCGGCGTTCGCGCATCTC	3000
Qy	3001	CGACGCCATCAACAAAGGGGTACAGTCTTCGGGGGCGCTTCGGAAGCTCTTCCTCGC	3060
Db	3001	CGACGCCATCAACAAAGGGGTACAGTCTTCGGGGGCGCTTCGGAAGCTCTTCCTCGC	3060
Qy	3061	CTTTCGGGTTCATCGTCCACCTCTACCGTTTCCTCAAGGGGCTCATGGGGCGCAGAAC	3120
Db	3061	CTTTCGGGTTCATCGTCCACCTCTACCGTTTCCTCAAGGGGCTCATGGGGCGCAGAAC	3120
Qy	3121	GACGCCCATCGTTGTTGTCATCTGGTCCATTTCTGCTGGGCTTCATCTTCCTGCTCTG	3180
Db	3121	GACGCCCATCGTTGTTGTCATCTGGTCCATTTCTGCTGGGCTTCATCTTCCTGCTCTG	3180
Qy	3181	GGTCAGGATCGAACCTTTTCATGTCAGGACCAAGGGCGCGGACGTTCAGGAGTTCGCA	3240
Db	3181	GGTCAGGATCGAACCTTTTCATGTCAGGACCAAGGGCGCGGACGTTCAGGAGTTCGCA	3240

[illegible]

SOURCE	Oryza sativa (japonica cultivar-group)	
ORGANISM	Oryza sativa (japonica cultivar-group)	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
	Ehrhartoideae; Oryzoae; Oryza.	
REFERENCE	1	
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.	
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC	
	clone:P0418B08	
JOURNAL	Published Only in Database (2002)	
REFERENCE	2 (bases 1 to 165909)	
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (12-JUN-2002) Takuji Sasaki, National Institute of	
	Agrobiological Sciences, Rice Genome Research Program; Kannondai	
	2-1-2, Tsukuba, Ibaraki 305-8602, Japan	
	(E-mail:tsasaaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,	
	Tel:81-298-38-7441, Fax:81-298-38-7468)	
COMMENT	On Jul 27, 2004 this sequence version replaced gi:41152735.	
	Genes were predicted from the integrated results of the following:	
	GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH	
	(http://www.softberry.com/), GeneMark.hmm	
	(http://opal.biology.gatech.edu/GeneMark/), GlimmerM	
	(http://www.tigr.org/tdb/glimmerm/glmr.form.html), RiceHMM	
	(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor	
	(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4	
	(http://globin.cse.psu.edu/html/docs/sim4.html), gap2	
	(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The	
	genomic sequence was searched against NCBI NonRedundant Protein	
	database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA	
	sequence database at RGP or DDBJ. Protein homologies of the coding	
	regions were searched against NCBI NonRedundant Protein database	
	with BLASTP. ESTs represent the identified cDNA sequences using	
	BLASTN with the corresponding DDBJ accession no. and RGP clone ID.	
	Full-length cDNAs represent the identified cDNA sequences using	
	BLASTN with the corresponding DDBJ accession no.	
	A gene with identity or significant homology to a protein is	
	classified based on the protein name to indicate the homology level	
	such as same name, 'putative-' and '-like protein'. A gene without	
	significant homology to any protein but with full-length cDNA or	
	EST homology (covering almost the entire length of partial	
	sequence) is classified as an 'unknown' protein. A gene predicted	
	by two or more gene prediction programs is classified as a	
	'hypothetical' protein according to IRGSP standard. A gene	
	predicted by a single gene prediction program is also classified as a	
	probable 'hypothetical' protein and is included as a	
	miscellaneous feature of the sequence.	
	The orientation of the sequence is from SP6 to T7 of the PAC clone.	
	This sequence of P0418B08 clone has an overlap with OJ1740.D06	
	(DDBJ: AP005579) clone at 5' end and with OJ1299.All (DDBJ:	
	AP005568) clone at 3' end. Detailed information on overlap and	
	assembly quality together with annotation of this entry is	
	available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.	
FEATURES	Location/Qualifiers	
source	1..165909	
	/organism="Oryza sativa (japonica cultivar-group)"	
	/mol_type="genomic DNA"	
	/cultivar="Nipponbare"	
	/db_xref="taxon:39947"	
	/chromosome="9"	
	/clones="P0418B08"	
	/complement(5396..5638)	
gene	/gene="P0418B08.1"	
	/complement(5396..5638)	
misc_feature	/gene="P0418B08.1"	
	/note="hypothetical ORF	
	predicted by GlimmerM	
	this category is not included in IRGSP standard"	
gene	5856..10431	
	/gene="P0418B08.2"	
mRNA	join(5856..5221,6327..6562,6658..6754,6846..6972,	
	7059..7671,7761..8024,8107..8319,8442..8649,8741..8937,	
	9019..9372,9476..10431)	
	/gene="P0418B08.2"	
	/notes="supported by full-length cDNA(s): AK1211170"	
	join(5951..6221,6327..6562,6658..6754,6846..6972,	
	7059..7671,7761..8024,8107..8319,8442..8649,8741..8937,	
	9019..9372,9476..10063)	
	/gene="P0418B08.2"	
	/notes="contains EST(s): C72208(EI220), AU082309(EI220)	
	contains full-length cDNA(s): AK121170"	
	/codon_start=1	
	/product="putative cellulose synthase"	
	/protein_id="BAD33412.1"	
	/db_xref="GI:50725884"	
	/translation="MEASAGLVAGSHNRELVLIRGHEPKPLRALSGQVCEICGDEV	
	GRVDDGLFVAGNECGFPVPCRCYEREGTQNCQCKTRYKRLGSGPRVPGDEDE	
	DIDDLHEFNIDDEKQKQDQDQWQNSHITAMLHGMYSYGRGDDGNSPTLPP	
	IITGARSVPVSGEFPTISNGHGGEFSSLLHRIHPYFVSEPGSAKWEKKEVSKRM	
	DWKSQGVAGGAPDDYDADVPLNDEAROLSRKVSIASSKVNPKYRNVILRLVV	
	LQFFLYRLHPVDAIPLWLTISICEIWFVSMWILDQPKWYPIIDRETVLDRLSRLV	
	EREGESLISAVDLFVSTDPLKEPLVTANTVLSILADYDVKVSCYSDGASML	
	TESLSEATFAFKWPFCKKFSIEPRAPEFYFSQVDYLKOKHNFVYERRMKKE	
	YEFKVRINALVAKAQPAEGWIMKDGTPWGNTRDHPGMIQVFLHGSHGHTGEG	
	ELPLVYVSREKRPQGHKKAGANALIRVSAVLTNAPMLNLDCHDHYINNSKATRE	
	AMCFMLDPQGVKCVQFPQRPDGIHDVDRVANRTVPFDINMKGLDQFGQVYVGT	
	GVVERQALYGVNPKGPKRKMVTCDCPCFCRKRKKGKGLPRVADGMDSDK	
	EMLMSQNFKEFGQSAAFVTSTLMEGGVPSPPSAPALLKEAIIHVISCYEKTDWG	
	BLGWITYGSDITDILGFKWCHKRSVYVCMKRAAPKGSAPINLSDRLNQVLRWALG	
	SVEIFSRHSPLLYKNGNLKWLERSYINTIYPTSLPLLAYCTLPVACLLTKGF	
	IMPTISTFASLFFIALIFATGILEMRWSVSIEMWRNEQVIGWISAHLFVAV	
	QGLKVLAGIDNTVTSKATGDEDEFAELVAFKWTLLIPTTLLILNIIGVAGV	
	SDAINNGSAMGPLGKLPFAFWIVLHYPFLKGLMGRQNRTPITIVINSVLLASIFS	
	LWLRDPTTIIKARGPDVQGINC"	
gene	complement(10866..17118)	
mRNA	/gene="P0418B08.3"	
	/complement(join(10866..11328,11488..11595,12606..12695,	
	12806..12922,13597..13716,14119..15432,16817..17118))	
	/gene="P0418B08.3"	
	/notes="supported by full-length cDNA(s): AK100679"	
	complement(join(11263..11328,11488..11595,12606..12695,	
	12806..12922,13597..13716,14119..15357))	
	/gene="P0418B08.3"	
	/notes="contains full-length cDNA(s): AK100679"	
	/codon_start=1	
	/product="streptococcal hemagglutinin-like protein"	
	/protein_id="BAD33413.1"	
	/db_xref="GI:50725885"	
	/translation="MDVLKSDVKTKILNETLPLPLVPSEKHNASPVNRGRDVASRYK	
	NGLSAHSATARTCTSPSGRTSANECTPEPKRAQSDRRRSTPSSRYSTSTPAS	
	RSVTPVRNTVTEGHKSRRITSTRNTDGLWPMRNLSSFSQSESVTPGNKDKVVP	
	GLSDQTKGASVIAERKGRPLRRKNIGECENAQPSDQPRRVIEQHRVAMQSGRVA	
	SNILSIDMSKAGRSVPTNISRGVSPKRLTASEGTGKFNKSLDEPARRLAIHAG	
	GRDDKVDSSCHAYSQSQRCKSVSRPRAVTLFVPLHRSSSPSKASVTSISRSFQ	
	SPSRTPSPSRQSAGSIQSGVASPIISVMVDKKNKNSQIENIHQHLGSYRFL	
	OMIFNVAEDTMSFOKVTAEIITYNVLRNTSLRDVNMRRIMVOCIQQLKHLGIL	
	KEQIDYLEQWPALEKENSISLPHATEALKASTLRLLPVTSGAKADVVALKVASVDI	
	MGLGSAPRCMLPKVEDRILYSELVSIARQEKMLDECRELLAMAALQVQSSSLRT	
	HITQURPQIAHMI"	
gene	18866..19216	
	/gene="P0418B08.4"	
misc_feature	18866..19216	
	/gene="P0418B08.4"	
	/note="hypothetical ORF	
	predicted by GlimmerM	
	this category is not included in IRGSP standard"	
gene	25510..27251	
	/gene="P0418B08.5"	
misc_feature	25510..27251	
	/gene="P0418B08.5"	
	/notes="probably inactive due to including stop codon(s) in	
	CDS	
	pseudogene, transposase"	
gene	/pseudo	
	complement(29005..36438)	
gene	/gene="P0418B08.6"	

gene
FACKQLEDGRTADYNQKESTLHLVLRGGTMIKVKTLTGKEIIDEPTDIDRI
complement (2464..8124)
/gene="OJ1740_D06.2"
mRNA
complement (join(2464..3329,3566..4732,5547..5616,
7723..8124))
/gene="OJ1740_D06.2"
/notes="supported by full-length cDNA (s): AK065475"
complement (join(2772..3329,3566..4732,5547..5616,
7723..7784))
CDS
/gene="OJ1740_D06.2"
/notes="containing EST(s): AU092173(C11159), D22463(C11159)
contains full-length cDNA(s): AK065475"
/codon_start=1
/product="putative RPT2"
/protein_id="BAD33627.1"
/db_xref="GI:50726106"
/translation="MKFMKIGSKPDAPQSDGADVVRVVISPLATDVIVHSEVKFYLHK
PFLIKSSKLQRLVIRKATRGTEDEVHIDGPGVTAPEICAKFCYGNVTLSPHNVA
ARCAEYTEMEDVDGNLFIKIDVFNSSILRSWKDSIIIVLOSTKALLPWSBELKVI
GRCIDAIASKTSDPANVTSYSHRKGMSCTEIVSGTSTIAPKDWVYEDLCELD
DLKRVVAVKSGRSPELIGALKAYAVRPLDSYDALVAEDYMRNOCIVETIIV
LLPSDKTSGCSGFLKLLKVAIIVGAGOHVKEELMERISFOLHKA SVKDLALPAASP
SDGADHVKLVNLRVPAVTAMSHNGGFVEKSDDKMIELNFQESTLALGELYDGYL
SEVASPDLSTFVELATVAPPAARPHVDSIDSAVDAYKHEPNISKADKKICGLI
DVKKLSDASHMATQNDRLPLRVQVLPQQLRAGSSNALTDGGGHTCAKPIMKD
QSDICERRIPRHNSLNKQATLSAREVHRKSEHGRGSRNFKDQGLFLQSRSR
IPDKIWSKQGGKSGSETSGSPPLSAKPADVPKPSPLPLRNRYSVS"
/gene="OJ1740_D06.3"
complement (join(10055..10068,13432..13638,13766..13899))
/gene="OJ1740_D06.3"
/notes="hypothesized ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
14864..15751
/gene="OJ1740_D06.4"
14864..15751
/gene="OJ1740_D06.4"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative HCWP repeat containing protein"
pseudo
16561..17118
/gene="OJ1740_D06.5"
16561..17118
/gene="OJ1740_D06.5"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative HCWP repeat containing protein"
pseudo
complement (17924..18233)
/gene="OJ1740_D06.6"
complement (join(<17924..17955,18104..>18233))
/gene="OJ1740_D06.6"
/notes="start and end point are not identified"
complement (join(17924..17955,18104..18233))
/gene="OJ1740_D06.6"
/notes="similar to Oryza sativa chromosome 1, P0480E02.21"
/codon_start=1
/product="hypothesized protein"
/protein_id="BAD33628.1"
/db_xref="GI:50726107"
/translation="MKMKKIVDDDKRNHRDEEDVDGDPANSSVATNDNDKARQHLHN
ASLGELMK"
gene
join(18381..18410,18539..18577,19649..19732)
/gene="OJ1740_D06.7"
misc_feature
join(18381..18410,18539..18577,19649..19732)
/gene="OJ1740_D06.7"
/notes="hypothesized ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
25476..26366
/gene="OJ1740_D06.8"
25476..26366
misc_feature

gene
/gene="OJ1740_D06.8"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, TNP2-like transposon protein"
pseudo
26747..28873
/gene="OJ1740_D06.9"
<26747..28873
/genes="OJ1740_D06.9"
/notes="supported by full-length cDNA (s): AK065973"
27353..28669
/gene="OJ1740_D06.9"
/notes="putative Rim2 protein"
29731..32622
/gene="OJ1740_D06.10"
29731..32622
/gene="OJ1740_D06.10"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative hydroxyproline-rich glycoprotein"
pseudo
join(36569..36654,36950..37116,37456..37510,38073..38244,
38373..38573)
/genes="OJ1740_D06.11"
join(36569..36654,36950..37116,37456..37510,38073..38244,
38373..38573)
/genes="OJ1740_D06.11"
/notes="hypothesized ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement (41633..42004)
/gene="OJ1740_D06.12"
complement (<41633..>42004)
/genes="OJ1740_D06.12"
complement (41633..42004)
/notes="start and end point are not identified"
/gene="OJ1740_D06.12"
/notes="predicted by GlimmerM etc."
/codon_start=1
/product="hypothesized protein"
/protein_id="BAD33629.1"
/db_xref="GI:50726108"
/translation="MTDEWPRVVRVRAELAVTGRGRPHVAAVVSLSYSGNHWSREIDDP
LITLFFGLSLSCSGSREARETARDDDHCCSCFVRVAGARSVSVIRLSQASRLPLH
AKSSPNSMRMRETVARTEHY"
42745..42927
/gene="OJ1740_D06.13"
Query Match 2.3%; Score 79; DB 8; Length 187410;
Best Local Similarity 100.0%; Pred. No. 6.1e-30; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 0;
Qy 817 GAACGACGAGCGGAGCGAGCGCGCTGTCGAGGAAGGTGTGATCGCGTCGAGCAAGGTGAA 876
Db 121846 GAACGACGAGCGGAGCGAGCGCGCTGTCGAGGAAGGTGTGATCGCGTCGAGCAAGGTGAA 121905
Qy 877 CCCGTACCGGATCGGTGATC 895
Db 121906 CCCGTACCGGATCGGTGATC 121924
RESULT 5
AY483156
LOCUS
DEFINITION
Hordeum vulgare putative cellulose synthase catalytic subunit
(CesA8) mRNA, partial cds.
ACCESSION
AY483156
VERSION
AY483156.1
KEYWORDS
GI:39726036
ORGANISM
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE
1 (bases 1 to 1255)

AUTHORS Burton,R.A., Shirley,N.J., King,B.J., Harvey,A.J. and Fincher,G.B.
TITLE The Cesa Gene Family of Barley. Quantitative Analysis of
Transcripts Reveals Two Groups of Co-Expressed Genes
JOURNAL Plant Physiol. 134 (1), 224-236 (2004)
PUBMED 14701917
REFERENCE 2 (bases 1 to 1255)
AUTHORS Burton,R.A., Shirley,N.J., King,B.J., Harvey,A.J. and Fincher,G.B.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2003) School of Agriculture & Wine, University of
Adelaide, PMB 1, Glen Osmond, SA 5064, Australia
FEATURES
source Location/Qualifiers
1. .1255 /organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
gene 1. .1255 /gene="Cesa8"
CDS <1..1090 /gene="Cesa8"
/codon_start=2
/product="putative cellulose synthase catalytic subunit"
/protein_id="AA029968.1"
/db_xref="GI:39726037"
/translations="GRVGSAAFTVTSMEEGGVPPSSPAALLKEAIHIVISCGYEDKT
DMLEIGWIGSTEDILTGFRMCHGRWSIYCMPLAAPKGSAPINLSDRLNQVLRW
ALGSVRIFFSRHSPLLGYKGNLKWLERPAYINTIYPTSLPLLAYCTLPVAVCLLT
GKFIPISTFASLFFISLIFISFATGILELRWSGYSIEWCRNEOLWIGVWAHLF
AVIQGLLWAGIDTKFTVTSKATGYEDDEFARLYAFKWTLLIPOTTLILNIGLV
AGISDAINNGYQSWGLPKLFPAPFVIVHLFPFLKGFMRQNRFTTIVIIWSVLLAS
MFSLWVRIDPFTVKAGPDVKKQGCNC"
ORIGIN
Query Match 1.7%; Score 59; DB 8; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3.1e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2243 CTCCTCAAGAGGCCATCCATGTCATCAGCTCGGCTACGAGCAAGACCGACTGGG 2301
Db 83 CTCCTCAAGAGGCCATCCATGTCATCAGCTCGGCTACGAGCAAGACCGACTGGG 141
RESULT 6
AY372244 3470 bp mRNA linear PLN 20-JAN-2004
LOCUS Zea mays cellulose synthase catalytic subunit 10 (Cesa10) mRNA,
DEFINITION complete cds.
ACCESSION AY372244
VERSION AY372244.1 GI:38532099
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3470)
AUTHORS Dhugga,K.S., Barreiro,R., Appenzeller,L., Wang,H., Niu,X.,
Carrigan,L. and Tones,D.
TITLE Cellulose formation and its role in determining stalk strength in
maize
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3470)
AUTHORS Dhugga,K.S.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-2003) Trait and Technology, Pioneer Hi-Bred
International, Inc., a DuPont company, 7300 NW 62nd Avenue,
Johnston, IA 50131, USA
FEATURES
source Location/Qualifiers
1. .3470 /organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
gene 1. .3470 /gene="Cesa10"

CDS
29. .3265
/gene="Cesa10"
/note="Cesa10; ZmCesa10"
/codon_start=1
/product="cellulose synthase catalytic subunit 10"
/protein_id="AAR23310.1"
/db_xref="GI:38532100"
/translation="MDAGSVTGGLAAGSHMRDELHVMAREEPNAKRVASADVKTCRV
ADEVTRDGDQPFVACAECPVPCVYEVSEGTQCPQCNTRVKRQKGPVSGD
EEGPEMDDFEFPAKSPKPEVADVYSENGEHPAKWRTGQTLSSFTGSVAG
KDLBAERMEGSMENKDRIDKWKTKQEKRGKLNHDDSDDDDKNEDEYMLAARQPL
WRKVPSPMINPYRIVIVLRVLVLCFFKFRITTPATDAVPLMLASVIELWFAFSW
ILDQLKPAPVTRITYRLALRYDREGACRLSPIDFVFTVPLDLPKEIPIITANTVL
SILADYVPDRVSCVSDRRAKMLLFDALESABFARWVPFCKFAVPFAPFEPFYS
OKIAYDKVQVPTFKVRRAMREYEEFKRINALVAKVAKQPEEGWQDGPWPGN
NTRDHPGMIOTVLNGQALDVEGHELPLVIVYSREKRCYGNHHKAGANALVRVSAV
LNTRAPFILNLDCHYVNSKAVREACFLMDPQLGKLCYVQFPQRFDCIGDRHRYAN
RVNVFDINMKDGIQGVYVTCVFNQALYGDPPREKRPMTCDPWSWCC
CCFGGKRGKARKDKGDGEEPRGLLGYFKRKKDKLGGSVAGSKGGGLYKK
HORAFLEEEIEREGYDELSRLMSQSKSFEKRFQSPVFIASLVEDGGLPQGA
DPAALKEAIHIVISCGYEKTEWKEIGVIYGSVTEIDILTGFRMCHGRWSIYCMPL
PAKGSAPINLSDRLNQVLRWALGSVVEIFMSRCHPLRAYGGRKLWLERPAYINTIY
PFTSLPLAYCTLPVAVCLLTGKFIPISTFASLFFISLIFISFATGILELRWSGYS
EDWRNEQFVWIGVSAHLFAVQGLKVLGGVDTSFTVTSKAGDEADFGDLYLFK
WTLIVPPTTLIIINNVGIVAGVDANNVGSWGPLFGKLFPSFWIVHLFPFLKGL
MGRQNRTPITVVLWSILLASIFSLVWVRIDPFIKAGPILKPCGVEC"
ORIGIN
Query Match 1.5%; Score 50; DB 8; Length 3470;
Best Local Similarity 100.0%; Pred. No. 1.7e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2771 AACGAGCAGTTCTGGGTCAATCGCGCGTGTGGCGCATCTCTTCGCCGT 2820
Db 2786 AACGAGCAGTTCTGGGTCAATCGCGCGTGTGGCGCATCTCTTCGCCGT 2835
RESULT 7
BC059699 814 bp mRNA linear VRT 28-SEP-2004
LOCUS Danio rerio chromosome 20 open reading frame 149, like, mRNA (cdna
DEFINITION BC059699
ACCESSION BC059699
VERSION BC059699.1 GI:37590413
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 814)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahney,J., Helton,E., Kettner,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Schnurch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932

REFERENCE 2 (bases 1 to 814)
 AUTHORS Director MGC Project.
 TITLE Direct Submission
 JOURNAL Submitted (01-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb-remail.nih.gov
 Tissue Procurement: Susan E. Brockerhoff, University of Washington
 cDNA Library Preparation: Susan E. Brockerhoff, University of Washington
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 136 Row: j Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 41152387.

FEATURES

source
 1..814
 Location/Qualifiers
 /organism="Danio rerio"
 /mol_type="mRNA"
 /strain="wild-type"
 /db_xref="taxon:7955"
 /clone="MGC:73395 IMAGE:4787352"
 /tissue_type="Eye, adult retina, 1-2 year old, mixed sex"
 /clone_lib="zebrafish adult retina cDNA"
 /lab_host="DH10B"
 /note="vector: pBluescript SK-"

gene

1..814
 /gene="c20orf1491"
 /note="synonyms: f62f01, zgc:73395"
 /db_xref="LocusID:336303"
 166..513
 /gene="c20orf1491"
 /codon_start=1
 /product="Unknown (protein for MGC:73395)"
 /protein_id="AAH59699.1"
 /db_xref="GI:37590414"
 /db_xref="LocusID:336303"
 /translation="MAAIPASGLVATHDYRRIGTSSSSCGSGSEYGEVPHHP
 GLPKQSGHWSFFFGKQLNPGMTVAEEAQKSGVSVTNGQVTCVAREMVMRQAS
 ESSDGGKSEAGNS"

CDS

Query Match 1.4%; Score 49; DB 5; Length 814;
 Best Local Similarity 100.0%; Pred. No. 6.2e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Qy 3395 TTAATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
 |||||
 Db 754 TTAATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 802
 |||||

RESULT 8

LOCUS HSM808426 4919 bp mRNA linear PRI 30-AUG-2003
 DEFINITION Homo sapiens mRNA; cDNA DKFZp686B20125 (from clone DKFZp686B20125).
 ACCESSION BX648278
 VERSION BX648278.1 GI:34367437
 KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Bloecker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
 CONSRMTM The German Human cDNA Consortium
 TITLE Direct Submission
 JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr. 1, D-85764 Neuherberg, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project
 This clone (DKFZp686B20125) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source
 1..4919
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686B20125"
 /tissue_type="human uterus"
 /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"
 /dev_stage="adult"
 4844

ORIGIN

Query Match 1.4%; Score 49; DB 9; Length 4919;
 Best Local Similarity 100.0%; Pred. No. 5.6e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3395 TTAATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
 |||||
 Db 4836 TTAATGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4884
 |||||

RESULT 9

LOCUS BC083654 1368 bp mRNA linear ROD 01-OCT-2004
 DEFINITION Rattus norvegicus cDNA clone MGC:94487 IMAGE:7135192, complete cds.
 ACCESSION BC083654
 VERSION BC083654.1 GI:53733462
 KEYWORDS MGC.

SOURCE

Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Atkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Platchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Carnicelli, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Abramson, R.D., Mullany, S.J., Bosak, S.A., Loquellano, N.A., Peters, G.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

/translation="MGKNSKLREVLQDLRENTETDHELOEWYKGLKDCPTGHLT
VDFPKLIYANFPYDASKAEHVFRTDNGDTDFRFLIALSVTSRGLKLEQKX
WAFSDLDGNGISSEMLEIYQAIYKVMSSWKMPEDSTPEKRTDKIFRQMDTNN
DGKLSUEEPIRGAKSDPSIVRLQLQCPSSASQ"

ORIGIN
Query Match 1.4% Score 48; DB 9; Length 1792;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443

Db 1736 TAAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1783

RESULT 11
AC141002/c
LOCUS
DEFINITION
Homo sapiens 3 BAC RP11-57206 (Roswell Park Cancer Institute Human
BAC library) complete sequence.

AC141002
AC141002.8 GI:29244589
HTG.
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 138981)

AUTHORS
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,P.R., Allen,C.,
Alsbrooks,S.L., Amaral,J., Bimonte,J., Bimonte,J., Bimonte,J.,
Barbieri,J., Benton,J., Bimonte,J., Bimonte,J., Bimonte,J.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Brown,D.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,H.J.,
Dunham,J., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Honsi,P., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kreshi,A., Landry,N., Leal,B., Lewis,D.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Morris,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,N., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokoko,S., Oguh,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojehokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S.,
Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E.,
Sonaite,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wlezyk,R., Woodson,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

REFERENCE
AUTHORS
JOURNAL

REFERENCE
AUTHORS
JOURNAL

REFERENCE
AUTHORS
JOURNAL

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht

ml.

FEATURES
source

Location/Qualifiers
1. .138981
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-57206"
complement(144..215)
/rpt_family="AluYd3a1"
complement(231..344)
/rpt_family="AluSq/x"
587..701
/rpt_family="L2"

repeat_region
repeat_region
repeat_region

```
repeat_region complement(702..1016)
repeat_region /rpt_family="AluY"
repeat_region 1017..1172
repeat_region /rpt_family="L2"
repeat_region 1173..1412
repeat_region /rpt_family="L1"
repeat_region 1413..1644
repeat_region /rpt_family="AluSq"
repeat_region 1645..1795
repeat_region /rpt_family="L1"
repeat_region 1796..1938
repeat_region /rpt_family="AluSx"
repeat_region 1939..2239
repeat_region /rpt_family="AluY"
repeat_region 2240..2415
repeat_region /rpt_family="AluSx"
repeat_region 2416..2532
repeat_region /rpt_family="L1"
repeat_region 2548..2842
repeat_region /rpt_family="AluY"
repeat_region 2888..3108
repeat_region /rpt_family="AluJo"
repeat_region 3111..3212
repeat_region /rpt_family="L1"
repeat_region 3284..3426
repeat_region /rpt_family="L1"
repeat_region complement(3460..3767)
repeat_region /rpt_family="AluSc"
repeat_region complement(3831..3861)
repeat_region /rpt_family="Alu"
repeat_region 3870..3898
repeat_region /rpt_family="AT rich"
repeat_region complement(3900..4188)
repeat_region /rpt_family="AluJb"
repeat_region 4220..4336
repeat_region /rpt_family="LIME"
repeat_region 4360..4525
repeat_region /rpt_family="LIME"
repeat_region 4526..4817
repeat_region /rpt_family="AluSx"
repeat_region 4818..5134
repeat_region /rpt_family="LIME"
repeat_region 5136..5338
repeat_region /rpt_family="AluSx"
repeat_region complement(5411..5676)
repeat_region /rpt_family="AluJo"
repeat_region 6447..6752
repeat_region /rpt_family="AluSx"
repeat_region 6758..7062
repeat_region /rpt_family="AluSq"
STS 7187..7310
/standard_name="RH93866"
repeat_region 7479..7597
repeat_region /rpt_family="FLAM_A"
repeat_region 8012..8297
repeat_region /rpt_family="AluJb"
repeat_region 9581..9654
repeat_region /rpt_family="L2"
repeat_region 9755..10198
repeat_region /rpt_family="L2"
repeat_region 10220..10270
repeat_region /rpt_family="(TA)n"
repeat_region complement(10273..10569)
repeat_region /rpt_family="AluSx"
repeat_region 10743..11201
repeat_region /rpt_family="L2"
repeat_region complement(11202..11516)
repeat_region /rpt_family="AluJb"
repeat_region 11517..11527
repeat_region /rpt_family="L2"
repeat_region complement(11528..11800)
Query Match 1.4%; Score 48; DB 9; Length 138981;

Best Local Similarity 100.0%; Pred. No. 1.6e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
DB 63816 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 63769

RESULT 12
CQ700614 CQ700614 431 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 45540 from Patent WO02070737.
ACCESSION CQ700614
VERSION CQ700614.1 GI:42261381
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 45540 12-SEP-2002;
ChondroGene Inc. (CA)
FEATURES
Location/Qualifiers
source 1..431
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 1.4%; Score 47; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.4e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
DB 337 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 383

RESULT 13
BC058737 BC058737 1409 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus RIKEN cDNA 1700020C11 gene, mRNA (cDNA clone MGC:68085
DEFINITION IMAGE:5400435), complete cds.
ACCESSION BC058737
VERSION BC058737.1 GI:37589295
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1409)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
```


Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1431)
Strausberg,R.
Direct Submission
Submitted (01-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: anadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAK Plate: 84 Row: i Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 8923916.
Location/Qualifiers
1. .1431
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:48766 IMAGE:5722236"
/tissue_type="Ovary, pooled from 3 adults"
/clone_lib="NIH MGC_125"
/lab_host="DH10B"
/note="Vector: pCMV-SPORTc"
1. .1431
/gene="KIAA1704"
/note="synonyms: AD029, ISR7, ba245H20.2"
/db_xref="LocusID:55425"
104. .1126
/gene="KIAA1704"
/codon_start=1
/product="KIAA1704"
/protein_id="AAH39586.1"
/db_xref="GI:24660420"
/db_xref="LocusID:55425"
/translation="MARDLIGPALPGFKVARGTAEDRPSVAGPALPNVKKSSSS
DSDSPEDSSSLVEQNGQESDGGPTARKOKRDDDDDDGFFGALPPGPKKQ
DSDPPRTIGPALPGFIKSTQKGRDDPGQETDSEDEDDIGMPFAKGFVNYV
TEFEKRAQRMKKLTKGDSDSKPIRESMTPELPPMKQFGLGRTIPRADDTSG
DSISWTDTPADREKAKETQAKRSKSKDEEHILSKDLAEQVSSYNESKRSSEL
MDIHHKLLKSKAEDKNKQERIPFRDKDLKVNRPDEAKKALIKKSRELNTRFSHG
KGNMFL"

FEATURES
source

gene

CDS

ORIGIN

Query Match 1.4%; Score 47; DB 9; Length 1431;
Best Local Similarity 100.0%; Pred. No. 6.9e-13;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3397 AAATGCCAAAAA 1431
Db 1384 AAATGCCAAAAA 1430

RESULT 16
LOCUS BC019205 1450 bp mRNA linear ROD 29-JUN-2004
DEFINITION Mus musculus RIKEN cDNA 1700020C11 gene, mRNA (CDNA clone MGC:29077
IMAGE:5051472), complete cds.
ACCESSION BC019205
VERSION BC019205.1 GI:17512505
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1450)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1450)
Strausberg,R.
Direct Submission
Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Grant,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Scantripo,S., Thomas,P.J., Touchman,J.W.,
Tsurgon,C., Vogt,J.I., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAK Plate: 39 Row: m Column: 2
This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 6678975.

FEATURES
source
Location/Qualifiers
1. .1450
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:23077 IMAGE:5051472"
/tissue_type="Liver, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_Li9"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .1450
/gene="1700020C11Rik"
/db_xref="LocusID:67900"
/db_xref="MGI:1916686"
160. .660
/gene="1700020C11Rik"
/codon_start=1
/product="RIKEN CDNA 2610507A21"
/protein_id="AAH19205.1"
/db_xref="GI:17512506"
/db_xref="MGI:1916686"
/translation="MSEQQQRGAERDLYRTWRYLGVANVGAFSLVPAVVWLS
YGVSSVYLAIDDKGKAGEVPSPEAGRNTRMALAVVDTFVWQSLASVAIPGFTINR
LCAASLYLGTMTWPTVRKWTITLGLLAIPVIHPIIDRSVDFDLSSLRKLYPSV
EKPSTP"

ORIGIN
Query Match 1.4%; Score 47; DB 10; Length 1450;
Best Local Similarity 100.0%; Pred. No. 6.9e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
Db 1400 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1446
|||||

RESULT 17
BC048832 1897 bp mRNA linear ROD 07-OCT-2003
LOCUS Mus musculus transducin-like enhancer of split 6, homolog of
Drosophila E(spl), mRNA (CDNA clone MGC:55024 IMAGE:4506029),
complete cds.
BC048832
VERSION BC048832.1 GI:28981430
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1897)
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalek,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL 22388257
MEDLINE 12477932
PUBMED 2 (bases 1 to 1897)
REFERENCE Straussberg,R.
AUTHORS Direct Submission
TITLE Submitted (14-MAR-2003) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: csapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stankirop,S., Thomas,P.J., Touchman,J.W.,
Taurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 99 Row: e Column: 7.
FEATURES
source
Location/Qualifiers
1. .1897
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:55024 IMAGE:4506029"
/tissue_type="Eye, retina, mouse strain C57Bl/6"
/clone_lib="NIH_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .1897
/gene="Tle6"
/note="synonyms: Gr96, 1810057E06Rik"
/db_xref="LocusID:114606"
/db_xref="MGI:2149593"
198. .1724
/gene="Tle6"
/codon_start=1
/product="Tle6 protein"
/protein_id="AAH48832.1"
/db_xref="GI:28981431"
/db_xref="LocusID:114606"
/db_xref="MGI:2149593"
/translation="MSGVGAEGSDWPCVSHSHEAPMGSAQTENSKEEDKQVPESA
LQHPKSTPGPOLPTRRRLFSDELQDPQVMDAEPQFCOGFLIOGLWELFMDSRQ
KNQEGGEDSDQSKDGLCDFKPEQPHRNSLSDADPFLIKSPSALLDYQEDV
SRQPEQESSGCRADKFLKSLGWSVESSCNQPSALWQLERFTVPQALQKRVLK
HQELLVVAVSSTFRHVFTCSQGIKVMNLVNVQAEEDPESHUKSCVQDNKVVLRIC
LLSNSRTLPAGGYNLPGVITWDLAASPLYEKCOLPCEGLSCQALANTKENMALGFT
DGTVRITWDLRTQEIKNLPGVTMDLAPSLYEKCOLPCEGLSCQALANTKENMALGFT
LFQSIIMSLAHSPTEDWLLGLANGQCLFNSRKRDQVLTVDTKDNTILGKFSFNGK
WASVGNGNFTVHSMPTGAKLFQVPEVGPVRCFDMTENGRLITGSRDCASVHIKY
" 534. .1694
/gene="Tle6"
/note="COG2319; Region: FOG: WD40 repeat [General function
prediction only]"
/db_xref="CDD:COG2319"
misc_feature
ORIGIN

```

Query Match      1.4%; Score 47; DB 10; Length 1897;
Best Local Similarity 100.0%; Pred. No. 6.8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
    |||
Db 1848 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1894
    |||

RESULT 18
BC043552
LOCUS BC043552 Homo sapiens, clone IMAGE:5174044, mRNA linear PRI 13-JAN-2003
DEFINITION Homo sapiens, clone IMAGE:5174044, mRNA.
ACCESSION BC043552
VERSION BC043552.1 GI:27694536
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE
1 (bases 1 to 2094)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcapsb-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Leticia Hsiao, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Santana Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saedi, Jacqueline
Schein, Duane Smalius, Michael Tsai, Natasja van den Bosch, Jill Vardy,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 87 Row: 1 Column: 19.
FEATURES
source
Location/Qualifiers
1..2094
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5174044"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH MGC_115"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"

ORIGIN
Query Match      1.4%; Score 47; DB 9; Length 2094;
Best Local Similarity 100.0%; Pred. No. 6.7e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3397 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
    |||
Db 2031 AAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2077
    |||

RESULT 19
BC018554
LOCUS BC018554 Mus musculus DNA segment, Chr 7, Brigham & Women's Genetics 0611
DEFINITION Mus musculus DNA segment, Chr 7, Brigham & Women's Genetics 0611

expressed, mRNA (cDNA clone MGC:19375 IMAGE:2645082), complete cds.
BC018554
VERSION BC018554.1 GI:17391320
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE
1 (bases 1 to 2513)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Loquellano, N.A., Peters, G.J.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalius, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2513)
Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcapsb-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lousseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 23 Row: a Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 26665872.
FEATURES
source
Location/Qualifiers
1..2513
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="MGC:19375 IMAGE:2645082"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MMV."
/clone_lib="NCI CGAP_Lu29"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
1..2513
/gene="D7Bwg0611e"

gene

```


/db_xref="LocusID:52857"
/db_xref="MGI:105490"
135..2078
/gene="D7Bwg0611e"
/codon_start=1
/product="DNA segment, Chr 7, Brigham & Women's Genetics
0611 expressed"
/protein_id="AAH18554.1"
/db_xref="GI:17391321"
/db_xref="LocusID:52857"
/db_xref="MGI:105490"
/translation="MQSWYMLCTYKORNEDEPKLPSKLPSEARLIVDYSCALOREI
LQGRILSENWICFYSNFWETIISQLEVCVKKKEKAKLIPNAIQTSEKX
FFTSFGARDRCFLIFRLQWQNALEKTLSPRELWHLVHCQVSEGLTSEDEDYVCLP
QLNGLGSPGDMIALSDISPGAADHSQSPVGSRRGRTVNLRASSADADHGA
EDKEQDSDQADASSQTVTPVAPFLSEPTPPDQPTSSGLPLDLLSREELLTDS
SSSTGEEDLADALLDLSGRLLINSVFHMGARLQOMLFSDSPFLOQKFTDVT
LSPWSDSKCHQRLVLTPIPIQNLGPKSASVETOTLFRGPOAGCVVDSEVLQ
GIPQDTPYTAHRCILGLARNKARLVSSIRYRKQPSLVKSLIEKNWSGIEDIF
HHLDELAKAEKLSLEGGKDTRGLLSGRRKRLPLSWRHRDGPQHPDPDPTQISM
HTSGLSRFSESPVDQGPAGIPFSAVLVLSIVLTVIALNALFLYRLMSLERTATFP
ESWHSALAKGKFPQTATEWAEIILALQKHFHSEVHKWRQILRASVELLDKMFSLK
LHQGITVPDPLDTPQPDGFP"

ORIGIN

Query Match 1.4%; Score 47; DB 10; Length 2513;
Best Local Similarity 100.0%; Pred. No. 6.7e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3397 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 2437 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2483

RESULT 20
HSU31501 2902 bp mRNA linear PRI 30-NOV-1995
LOCUS Human fragile X mental retardation syndrome related protein (FXR2)
DEFINITION mRNA, complete cds.
ACCESSION U31501
VERSION U31501.1 GI:1098636
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2902)
AUTHORS Zhang,Y., O'Connor,J.P., Siomi,M.C., Srinivasan,S., Dutra,A., Nussbaum,R.L. and Dreyfuss,G.
TITLE The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2
JOURNAL EMBO J. 14 (21), 5358-5366 (1995)
MEDLINE 96080171
PUBMED 7489725

REFERENCE 2 (bases 1 to 2902)
AUTHORS Zhang,Y. and Dreyfuss,G.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-1995) Yan Zhang, Biochemistry and Biophysics, Howard Hughes Medical Institute, University of Pennsylvania, Clinical Research Building, Room 330, 422 Curie Boulevard, Philadelphia, PA 19104-6148, USA

FEATURES
source
1..2902
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p31.1"
/tissue type="brain"
/dev stage="fetus"
1..2902
/gene="FXR2"
228..2249

/gene="FXR2"
/codon_start=1
/product="fragile X mental retardation syndrome related protein"
/protein_id="AAC50292.1"
/db_xref="GI:1098637"
/translation="MGGLASGDVEPLVPEVRSGNGAFYKGFVKVDHDSVTTFPEN
NQSERQIPFGDVELPPADYNKEITGEDEVYSRANEQPCQWLAARVMKGDYF
VIEAACDATTNEIVTLERLPVNPENLATKSGFKVTMAVPEDLREACSNVHKEF
KKALGANCFINLITNSFLSTTEAPVWRASLLGDHFRSLRKLILMSNEEATKH
LETSKQAAASQLEFSEDSQVPRNLGKVLGKQVLEIQLVDGSGVVRVEGDNDX
KNPREGQVFFIVGTRENTSNAQALLEHLSYLQEVLEQLERLQIDDELRQIGLGF
RPPSGSGSGSKAGYSTDDESSSSSHATKTYGGYGGRRGRTGGTGPAGPSSDVS
TASSTSEKKEPNRAGPDRDPTTGGESRRRTTGGRGRTGPAPRTSYNSSIS
SVLKDPSNPYSLLDTSPPVDSPEPPEPPASARRRRRRRTDEDTYNDGGLS
DGNMTNGLEDESRRPQRRNRSSRRNRNRTDGSISGDROPVTADVISAESQSRQ
SAPLERTKPSDESLSGQKSDSVSLPKGSPSENGELSLAPLELGSMVNGVS"

ORIGIN

Query Match 1.4%; Score 47; DB 9; Length 2902;
Best Local Similarity 100.0%; Pred. No. 6.6e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3397 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 2843 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2889

RESULT 21
BC054839 3099 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus armadillo repeat containing, X-linked 2, mRNA (cdna)
DEFINITION clone MGC:62562 IMAGE:6490133, complete cds.
ACCESSION BC054839
VERSION BC054839.1 GI:32766563
KEYWORDS MGC.
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)
REFERENCE 1 (bases 1 to 3099)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smillius,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 3099)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 467)
AUTHORS
Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
TITLE
The mosaic structure of variation in the laboratory mouse genome
JOURNAL
Nature 420 (6915), 574-578 (2002)
MEDLINE
22354684
PUBMED
12466852
COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 467
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGCv3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES
source
1..467
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C3H/HeJ"
/db_xref="taxon:10090"
/map="15 19-364 56751392-56751733"
/clone_lib="C3H/HeJ"
<1..>467

STTS
ORIGIN
Query Match 1.3%; Score 46; DB 11; Length 467;
Best Local Similarity 100.0%; Pred. No. 2.5e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAA 3443
|||||
Db 88 AATGGCAA 133
|||||

RESULT 24
BC043523
LOCUS
BC043523 Homo sapiens, clone IMAGE:5166551, mRNA.
DEFINITION
BC043523 Homo sapiens, clone IMAGE:5166551, mRNA.
ACCESSION
BC043523.1 GI:27694178
VERSION
BC043523.1
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 682)
AUTHORS
Strausberg,R.
TITLE
Direct Submission
JOURNAL
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT
Contact: MGC help desk
Email: cgapob@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
cDNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywninski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saesdi, Jacqueline
Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 87 Row: f Column: 4.

FEATURES
source
1..682
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5166551"
/issue_type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORTc"

ORIGIN
Query Match 1.3%; Score 46; DB 9; Length 682;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAA 3443
|||||
Db 594 AATGGCAA 639
|||||

RESULT 25
BC053171
LOCUS
BC053171 Danio rerio zgc:63964, mRNA (cDNA clone MGC:63964.6791988),
complete cds.
ACCESSION
BC053171
VERSION
BC053171.1 GI:31418813
KEYWORDS
MGC.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 1412)
AUTHORS
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toehiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abrahamson,R.D., Mullahy,S.J., Bosak,S.A., McSwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalys,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

REFERENCE	2 (bases 1 to 1412)	
AUTHORS	Strausberg, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Leonard I. Zon, M.D. cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.	
FEATURES	source	
	1. .1412	
	/organism="Danio rerio"	
	/mol_type="mRNA"	
	/db_xref="taxon:7955"	
	/clone="MGC:63964 IMAGE:6791988"	
	/tissue_type="Kidney, zebrafish"	
	/clone_lib="NCI CGAP_ZK1d1"	
	/lab_host="DH10B"	
	/note="Vector: pCMV-SPORT6.1"	
gene	1. .1412	
	/gene="zgc:63964"	
	/note="synonym: MGC63964"	
	/db_xref="LocusID:393323"	
CDS	196..1275	
	/gene="zgc:63964"	
	/codon_start=1	
	/product="hypothetical protein MGC63964"	
	/protein_id="AAH53171.1"	
	/db_xref="GI:31418814"	
	/db_xref="LocusID:393323"	
	/translation="MPAAVVDIKYKNDVLRFTKNAALPIIHPYNEVVVKVHAAGLNPI DISMRGIGTAATMAKMDPLNLSQSGGEFFLLIGRDVSGEIMECGLDVKYFKPGDQVM AATPPWQSGSLAEFPVVSNGNEVSHKSLRDEAASIPYVAATAWSAIVNTGGKDN SAKKRVILGGSGGVGTFAIQMKVKAHGVTVTCQNAERLVRDLGADDDVVDYTAGPV EKOLKNLEKFDLIDSGTEKVALDLKPWCAKFTVLTITPFLQNTDRILGLADGWM QSAVTVCCKVKNLRKGVHWGPFAPSGSALDEVSEWVDAKVRVVEVFSFAQVP EAFQKVPQHGARGKTIVSIMEDQKE"	
ORIGIN		
	Query Match 1.3%; Score 46; DB 5; Length 1412;	
	Best Local Similarity 100.0%; Pred. No. 2.3e-12;	
	Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	3398 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443	
Db	1363 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1408	
RESULT 26		
BC017176		
LOCUS	BC017176 1415 bp mRNA linear PRI 29-JUN-2004	
DEFINITION	Homo sapiens inositol(myo)-1(or 4)-monophosphatase 2, mRNA (cDNA	
	clone MGC:795 IMAGE:3345789), complete cds.	
ACCESSION	BC017176	
VERSION	BC017176.1 GI:16877898	
KEYWORDS	MGC.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 1415)	
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carlinici, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.	
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
PUBMED	12477932	
REFERENCE	2 (bases 1 to 1415)	
AUTHORS	Strausberg, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTF cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nhgri.nih.gov Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.	
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 5 Row: c Column: 4 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657235.	
FEATURES	Location/Qualifiers	
source	1. .1415	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="MGC:795 IMAGE:3345789"	
	/tissue_type="Skin, melanotic melanoma."	
	/clone_lib="NIH_MGC_20"	
	/lab_host="DH10B-R"	
	/note="Vector: pOTB7"	
	1. .1415	
	/gene="IMPA2"	
gene		
	/db_xref="LocusID:3613"	
	/db_xref="MIM:605922"	

CDS	83..949 /genes="IMPA2" /codon_start=1 /product="inositol(myo)-1(or 4)-monophosphatase 2" /protein_id="AAH17176.1" /db_xref="GI:116877899" /db_xref="LOCUSID:3613" /db_xref="MIM:605922" /translations="MKPSGEDQALAGPWEECFQAAVOLALFAGQIIRKALTEEKRV STKTSADLVTEHDIVEDLIISLRERPFPHFIAEBAASGAKGVLTHTSPWIIDP IDTCNFRHFPVAVSIGPAVQOELEFGVIVHCTERLYTGRGRGAPCNGORLVS GETDLSKALVLTIGKPRDPATKLFISNWERLLHAKAGVRVIGSGTALCHLASGA ADAYYQGLHCWDLAAATVIIRAGGIVIDTSGGPLDLWACRVVAASTREMAWLIAQA LQTYNGRDEK"	
ORIGIN	Query Match 1.3%; Score 46; DB 9; Length 1415; Best Local Similarity 100.0%; Pred. No. 2.3e-12; Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 3398 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443 DB 1363 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1408 	
RESULT 27	BC019066 1705 bp mRNA linear PRI 30-JUN-2004	
LOCUS	Homo sapiens solute carrier family 10 (sodium/bile acid	
DEFINITION	cotransporter family), member 4, mRNA (CDNA clone MGC:29802	
ACCESSION	IMAGE:4897519), complete cds.	
VERSION	BC019066	
KEYWORDS	BC019066.2 GI:33991607	
SOURCE	MGC.	
ORGANISM	Homo sapiens (human)	
REFERENCE	Homo sapiens	
AUTHORS	Homo sapiens	
1 (bases 1 to 1705)	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D.,		
Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,		
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,		
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,		
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Casavant,T.L.,		
Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,		
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,		
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,		
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,		
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,		
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,		
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,		
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,		
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,		
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E.,		
Schneerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
Generation and initial analysis of more than 15,000 full-length		
human and mouse cDNA sequences		
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
12477932		
2 (bases 1 to 1705)		
Strausberg,R.		
Direct Submission		
Submitted (07-DEC-2001) National Institutes of Health, Mammalian		
Gene Collection (MGC), Cancer Genomics Office, National Cancer		
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
USA		
NIH-MGC Project URL: http://mgc.nci.nih.gov		
On Aug 20, 2003 this sequence version replaced gi:17512161.		
Contact: MGC help desk		
Email: cgaps@mail.nih.gov		
Tissue Procurement: ATCC		
CDNA Library Preparation: Rubin Laboratory		
CDNA Library Arrayed by: The I.M.A.G.E.B. Consortium (LLNL)		
DNA Sequencing by: National Institutes of Health Intramural		
Sequencing Center (NISC),		
Gaithersburg, Maryland;		
Web site: http://www.nisc.nih.gov/		
Contact: nisc_mgc@nigri.nih.gov		
Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,		
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,		
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,		
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,		
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,		
McDowell,J., Pearson,K., Stantripop,S., Thomas,P.J., Touchman,J.W.,		
Tsurgou,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,		
Young,A., Zhang,L.-H. and Green,E.D.		
Clone distribution: MGC clone distribution information can be found		
through the I.M.A.G.E.B. Consortium/LLNL at: http://image.llnl.gov		
Series: IRAL Plate: 39 Row: j Column: 7		
This clone was selected for full length sequencing because it		
passed the following selection criteria: matched mRNA gi: 24308413.		
FEATURES	Location/Qualifiers	
source	1..1705	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="MGC:29802 IMAGE:4897519"	
	/tissue_type="Brain, neuroblastoma"	
	/clone_lib="NIH MGC_19"	
	/lab_host="DH10B-R_"	
	/note="Vector: pOTB7"	
gene	1..1705	
	/gene="SLC10A4"	
	/note="synonym: MGC29802"	
	/db_xref="LocusID:201780"	
CDS	131..1444	
	/genes="SLC10A4"	
	/codon_start=1	
	/product="solute carrier family 10 (sodium/bile acid	
	cotransporter family), member 4"	
	/protein_id="AAH19066.1"	
	/db_xref="GI:17512162"	
	/db_xref="LocusID:201780"	
	/translation="MDGNDNVLLFAPLRDNTVLAPDNASSLPGGTDLALAPASSAGP	
	HGSLSGPGSPFGSPPTPTPTTSGAGGAASGSPFPFPAPHALPFWDTPLN	
	HGLNVFGAALGCTMLGCTGVNVHFGHVRPVGALLAALCQGLPLPFLALLA	
	EKLDEAAVAVLLCGCCGNNLSLMDGDMNLSIIMTISLLALVMPCLMWI	
	YSWAWINTPVOLLPGTVTLT.CSTLPIGLGVETRYKYSRVADYIKVSLMSLIVT	
	LVLVFTGTMGPELLASIPAAVYVIALFMPLAGYASGYGLATLPHLPNCRTVCL	
	ETGSQNVQCTAILKLAFPPQFTGSMYMPFLLYALFQSAEAGIFVLIYKMYGSEMLHK	
	RDPLDEDEDTDISYKKLKEENADTSYGTVAENIIMETAQTSL"	
ORIGIN	Query Match 1.3%; Score 46; DB 9; Length 1705; Best Local Similarity 100.0%; Pred. No. 2.3e-12; Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 3398 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443 DB 1647 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1692 	
RESULT 28	AK026494 2258 bp mRNA linear PRI 13-SEP-2003	
LOCUS	Homo sapiens cDNA: FLJ22841 fis, clone KATA4844.	
DEFINITION	AK026494	
ACCESSION	AK026494.1 GI:10439367	
VERSION	oligo capping; fis (full insert sequence).	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

1
REFERENCE
AUTHORS
Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
Matsumura,K., Nakaajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hiraio,M.,
Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE
NEDO human cDNA sequencing project
JOURNAL
Unpublished
REFERENCE
AUTHORS
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE
Direct Submission
JOURNAL
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source
Location/Qualifiers
1. .2258
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KA1A484"
/tissue_type="ileal mucosa"
/clone_lib="kaia"
/note="cloning vector pME18SPL3"
ORIGIN
Query Match 1.3%; Score 46; DB 9; Length 2258;
Best Local Similarity 100.0%; Pred. No. 2.3e-12; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 0;
Oy 3398 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
Db 2201 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2246
|||||
RESULT 29
BC082653 2432 bp mRNA linear VRT 16-SEP-2004
LOCUS
Xenopus laevis cDNA clone MGC:81937 IMAGE:7009628, complete cds.
DEFINITION
ACCESSION BC082653
VERSION BC082653.1 GI:52138921
KEYWORDS
SOURCE MGC.
ORGANISM
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 2432)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
JOURNAL
Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 1245917
REFERENCE
AUTHORS
2 (bases 1 to 2432)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,P.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Hatschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettaman,M., Maman,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.C., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmezz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Schnurch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
12477932
PUBMED
3 (bases 1 to 2432)
REFERENCE
AUTHORS
Klein,S. and Gerhardt,D.S.
TITLE
Direct Submission
JOURNAL
Submitted (15-SEP-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
REMARK
COMMENT
Contact: XGC help desk
Email: gqabs-x@mail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhui,
Duaneh Saeedi, JR Santos, Angeliue Schnerch, Ursula Skalska,
Duane Smalusz, Jeff Scott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 162 Row: c Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
FEATURES
source
Location/Qualifiers
1. .2432
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:81937 IMAGE:7009628"
/tissue_type="Kidney, adult Xenopus"
/clone_lib="NICHD_XGC_Kidl"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
51. .1247
/codon_start=1
/product="Unknown (protein for MGC:81937)"
/protein_id="AAH82653.1"
/db_xref="GI:52138922"
/translation="MDGFQGGFKHFAIEHVGEPEQWKSMAVVPISLSTTFKQFG
PBGHAGDYDSRSGNPTNRNMEKVAALDGAAGKCYCLSYASGLAATINTHLKCGQIIC
TKDVGCTNRYFRVASEMGFKVAFVDCSLKSLAATITDTKLVTETPTNPLTVI
DIRKADIVHKKHDIILAVNDTFLSAYFORALGADI CMYSATKYNMGHSDVVMGLV
SVNCDKLYEKLFQNSLGAIPSPFCYKCNRLKTLHLMQHQFNALTVAKFLEAD
PRVDKLVLPGLSPHPOVELTKROCTGVSGWITFYIKGTLEHAKTEFLKSIKVPALAESL
GGYSLAEHPAINTHASVPEERAALNISTLRISLVGLEDEDLIEDLDQALGSAHP
DKNNH"
ORIGIN
Query Match 1.3%; Score 46; DB 5; Length 2432;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAAA 3443
|||||
Db 2363 AATGGCAAA 2408

RESULT 30

BC018294

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC018294 2500 bp mRNA linear ROD 29-JUN-2004
Mus musculus RIKEN cdna 1110061004 gene, mRNA (cdna clone MGC:19290
IMAGE:4021243), complete cds.

BC018294 BC018294.1 GI:17390693
MGC.
Mus musculus (house mouse)

Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2500)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Heiton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Schnerch,A., Schein,J.E., Jones,S.O. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 2500)
Strausberg,R.

Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-x@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loughsseg, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 24 Row: b Column: 11

This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction.

Location/Qualifiers
1. .2500
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"

FEATURES

source

/db_xref="taxon:10090"
/clone="MGC:19290 IMAGE:4021243"
/tissue type="Mammary tumor metastatized to lung."
MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR
enhancer."
/clone_lib="NCI CGAP_Lu30"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2500
/gene="1110061004Rik"
/db_xref="LocusID:97287"
/db_xref="MGI:1916075"
434. .2020
/gene="1110061004Rik"
/codon_start=1
/product="RIKEN CDNA 1110061004"
/protein_id="AAH18294.1"
/db_xref="GI:17390694"
/db_xref="LocusID:97287"
/db_xref="MGI:1916075"
/translation="MARCKRFPVCPVILFKGKHICRSATLAWGELYGRSGVNYLFSG
GADTWASTEDVTERDFVLRSGDTHLFDKVRGYDKLQYLSVKYICDLWENKKVKF
GNVTSSEKVDKAQRYANFTLASIPYGCFFKEVKRDYMAEGLIFNKQDVVDAPL
NIPNFTQSLNDMSQYQSNWLVQQTQNYLKLILFIMRDDSGLLVHICISGWDRTPL
FISLRLSLWADGLIHLSLKPAEILYLTVAIDWFLFGHMLVDRLSKGEEIFFCFNPL
KHITSEEFCLKTQRRKSLPTRDAGFTVEDICMLRHKORGSTTSLGSDFSLVLEHSPGA
VGSFSEYVELAPAGAPTOAWRKSHSSSPQMLMSRPQSEERLPSHGLTEAKSSS
SSSNVHSDNPFPMGSPLEVPKPSVDHPLPGSSLSLTDGFWOLVSGCCSIOBRPVLIH
TDSLLPFSFQDELPSLCILLSDRETLQEVRSFLAAYSSTVGLRAATPSPSGAIG
GULEQPARGVGLRGTSSTL"

gene

CDS

ORIGIN

Query Match 1.3%; Score 46; DB 10; Length 2500;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3398 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
Db 2431 AATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2476

Search completed: March 12, 2005, 05:53:47
Job time : 9668 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 23:55:25 ; Search time 1165 Seconds
(without alignments)
17494.995 Million cell updates/sec

Title: US-10-627-132-29
Perfect score: 3443
Sequence: 1 ctgcgtgcctgcctgcga.....aaaaaaaaaaaaaaaaaaaaa 3443

Scoring table: OLIGO NUC
Gapop=60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 25
Total number of hits satisfying chosen parameters: 105744

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

- Database : N_Geneseq_16Dec04.*
- 1: Geneseqn1980s.*
 - 2: Geneseqn1990s.*
 - 3: Geneseqn2000s.*
 - 4: Geneseqn2001as.*
 - 5: Geneseqn2001bs.*
 - 6: Geneseqn2002as.*
 - 7: Geneseqn2002bs.*
 - 8: Geneseqn2003as.*
 - 9: Geneseqn2003bs.*
 - 10: Geneseqn2003cs.*
 - 11: Geneseqn2003ds.*
 - 12: Geneseqn2004as.*
 - 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	1.5	1189	3	Aaz58264 Corn cell
2	48	1.4	907	9	Acd44873 Human sig
3	48	1.4	1795	6	Abq54275 Human ova
4	48	1.4	2218	12	Adq24226 Human sof
5	47	1.4	1826	4	Aah33312 Human col
6	45	1.3	189	3	Aaf22014 Human bre
7	45	1.3	480	9	Ach24208 Human adu
8	45	1.3	986	2	Aax51735 DNA encod
9	45	1.3	986	6	Abq92587 Human sec
10	45	1.3	1434	6	Abk33557 cDNA enco
11	45	1.3	1434	8	ACA68518 Novel hum
12	45	1.3	1434	9	ABT44247 Human PRO
13	45	1.3	1434	9	ABT44530 Human PRO
14	45	1.3	1434	9	ACD82197 Human sec
15	45	1.3	1434	9	ABT43903 Human mem
16	45	1.3	1434	9	ABD83533 Novel hum
17	45	1.3	1434	9	ADB80639 Novel hum
18	45	1.3	1434	9	ADB73180 Novel hum
19	45	1.3	1434	9	ADB78262 Novel hum
20	45	1.3	1434	10	ADB84910 Human PRO

21	45	1.3	1434	10	ADB78016	Novel hum
22	45	1.3	1434	10	ADB87082	Human PRO
23	45	1.3	1434	10	ADB84664	Human PRO
24	45	1.3	1434	10	ADB83779	Novel hum
25	45	1.3	1434	10	ADB72934	Novel hum
26	45	1.3	1434	10	ADC36772	Human PRO
27	45	1.3	1434	10	ADC21762	Human PRO
28	45	1.3	1434	10	ADC49793	Novel hum
29	45	1.3	1434	10	ADC48992	Novel hum
30	45	1.3	1434	10	ADC49509	Novel hum
31	45	1.3	1434	10	ADC47370	Novel hum
32	45	1.3	1434	10	ADC47115	Novel hum
33	45	1.3	1434	10	ADC77990	Novel hum
34	45	1.3	1434	10	ADC06225	Novel hum
35	45	1.3	1434	10	ADC7744	Novel hum
36	45	1.3	1434	10	ADD50707	Novel hum
37	45	1.3	1434	10	ADD50953	Novel hum
38	45	1.3	1434	10	ADD50434	Human PRO
39	45	1.3	1434	10	ADD50188	Human PRO
40	45	1.3	1434	10	ADD51199	Novel hum
41	45	1.3	1434	10	ADD89052	Novel hum
42	45	1.3	1434	10	ACA66862	Acad66862 cDNA enco
43	45	1.3	1434	10	ACD68614	Novel hum
44	45	1.3	1434	12	ADC48746	Novel hum
45	45	1.3	1434	12	ADC20917	Novel hum
46	45	1.3	1434	12	ADE05761	Human PRO
47	45	1.3	1434	12	ADD74990	Human PRO
48	45	1.3	1434	12	ADD75736	Novel hum
49	45	1.3	1434	12	ADD84968	Novel hum
50	45	1.3	1434	12	ADD86794	Novel hum
51	45	1.3	1434	12	ADE20671	Novel hum
52	45	1.3	1434	12	AD838968	Novel hum
53	45	1.3	1434	12	AD805515	Human PRO
54	45	1.3	1434	12	ADD73500	Human PRO
55	45	1.3	1434	12	ADD78340	Novel hum
56	45	1.3	1434	12	ADE21163	Novel hum
57	45	1.3	1434	12	ADD77278	Novel hum
58	45	1.3	1434	12	ADE20425	Novel hum
59	45	1.3	1434	12	ADD75490	Human PRO
60	45	1.3	1434	12	ADD74006	Human PRO
61	45	1.3	1434	12	ADD74252	Human PRO
62	45	1.3	1434	12	ADD75982	Novel hum
63	45	1.3	1434	12	ADD85474	Novel hum
64	45	1.3	1434	12	ADE05023	Human PRO
65	45	1.3	1434	12	ADD75236	Human PRO
66	45	1.3	1434	12	ADD76780	Novel hum
67	45	1.3	1434	12	ADD86548	Novel hum
68	45	1.3	1434	12	ADB78016	Novel hum
69	45	1.3	1434	12	ADD77524	Novel hum
70	45	1.3	1434	12	ADD77770	Novel hum
71	45	1.3	1434	12	ADD85228	Novel hum
72	45	1.3	1434	12	ADD73760	Human PRO
73	45	1.3	1434	12	ADD74498	Human PRO
74	45	1.3	1434	12	ADD77026	Novel hum
75	45	1.3	1434	12	ADD85720	Novel hum
76	45	1.3	1434	12	ADD85269	Human PRO
77	45	1.3	1434	12	ADD74744	Human PRO
78	45	1.3	1434	12	ADG05556	Novel hum
79	45	1.3	1434	12	ADG27110	Human PRO
80	45	1.3	1434	12	ADG11173	Novel hum
81	45	1.3	1434	12	ADG11952	Novel hum
82	45	1.3	1434	12	ADG94509	Novel hum
83	45	1.3	1434	12	ADG06605	Human PRO
84	45	1.3	1434	12	ADH38949	Novel hum
85	45	1.3	1434	12	ADG34039	Novel hum
86	45	1.3	1434	12	ADH33509	Human PRO
87	45	1.3	1434	12	ADH69603	Human PRO
88	45	1.3	1434	12	ADH29764	Novel hum
89	45	1.3	1434	12	ADM27161	Novel hum
90	45	1.3	1434	12	ADK66519	Human PRO
91	45	1.3	1576	5	AAD07711	Human sec
92	45	1.3	1772	3	AAF21809	Human bre
93	45	1.3	1947	8	AAD55845	Human nuc

c 94 45 1.3 1984 3 AAC59248 Aac59248 Human sec
95 45 1.3 2592 12 ADQ22667 Adq22667 Human sof
96 45 1.3 3118 3 AAF15989 Aaf15989 Human pro
97 45 1.3 3145 3 AAC59241 Aac59241 Human sec
98 45 1.3 3872 10 ADE79063 Ade79063 Human pro
99 45 1.3 4125 13 ADR44022 Adr44022 Human col
c 100 44 1.3 161 6 ABK54773 Abk54773 Human col

ALIGNMENTS

RESULT 1
AAZ58264
ID AAZ58264 standard; cDNA; 1189 BP.
XX AC AAZ58264;
XX
DT 08-MAY-2000 (first entry)
XX
DE Corn cellulose synthase cDNA crin.pk0135.e10.
XX
KW Corn; maize; cellulose synthase; transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 2..964
FT /*tag= a
FT /partial
XX
PN WO200004166-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015871.
XX
PR 14-JUL-1998; 98US-0092844P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JB, Miao G;
PI Rafalski JA, Thorpe CJ;
XX
XX WPI; 2000-182431/16.
DR P-PSDB; AAY58833.
XX
XX Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT probes for isolating cDNAs and genes encoding homologous proteins, for
PT producing transgenic plants.

XX
PS Claim 3; Page 42-43; 93pp; English.
XX
CC The present sequence is that of cDNA clone crin.pk0135.e10 encoding a
CC portion (see AAY58833) of corn cellulose synthase (CS). The cDNA clone
CC was isolated from a seedling root cDNA library on the basis of homology
CC to Arabidopsis and cotton CS sequences. The invention relates to isolated
CC nucleic acid fragments encoding plant CS and to CS polypeptides. It also
CC relates to the construction of a chimeric gene encoding all or a portion
CC of the CS, in sense or antisense orientation, where expression of the
CC gene results in altered levels of the CS in transformed host cells. The
CC host cells can be used to screen compounds for their ability to inhibit
CC CS activity. CS nucleic acids are also useful for producing transgenic
CC plants having altered levels of CS, and hence altered levels of fibre. CS
CC may also serve as a target for the development of novel herbicides
XX
SQ Sequence 1189 BP; 215 A; 374 C; 331 G; 269 T; 0 U; 0 Other;
Query Match 1.5%; Score 50; DB 3; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2771 AACGAGCAGTTCTGGGTGCATCGGCGGTGTCCGCGCATCTCTTCGCCGT 2820

Db 485 AACGAGCAGTTCTGGGTGCATCGGCGGTGTCCGCGCATCTCTTCGCCGT 534
RESULT 2
ACD44873
ID ACD44873 standard; cDNA; 907 BP.
XX
XX ACD44873;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human signal transduction pathway component HDPSE86 cDNA.
XX
KW Human; signal transduction pathway component; gene; ss; gene therapy;
KW neural disorder; immune system disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; multiple sclerosis; muscular disorder; sinusitis;
KW respiratory disease; nasal vestibulitis; nasal polyp; pulmonary disorder;
KW cardiovascular disorder; congenital heart defect; Ebstein's anomaly;
KW hypoplastic left heart syndrome; renal disorder; acute kidney failure;
KW end-stage renal disease; hyperproliferative disorder; Hodgkin's disease;
KW leukaemia; inflammatory disease; septic shock; bursitis; appendicitis;
KW allergy; asthma; blood related disorder; thrombosis; atherosclerosis;
KW myocardial infarction; endocrine disorder; Addison's disease; dysphagia;
KW corticosteroid deficiency; reproductive system disorder; dysmenorrhea;
KW testicular atrophy; gastrointestinal disorder; irritable bowel syndrome;
KW infectious disease; cancer; wound healing; epithelial cell proliferation.
XX
OS Homo sapiens.
XX
PN US2003036505-A1.
XX
PD 20-FEB-2003.
XX
PF 20-SEP-2001; 2001US-00955999.
XX
PR 25-SEP-2000; 2000US-0234997P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Ni J, Ruben SM, Rosen CA, Young PE, Rohrschneider LR;
XX
XX WPI; 2003-492122/46.
DR P-PSDB; ABO27187.
XX
XX New isolated nucleic acids encoding signal transduction pathway component
PT polypeptides, useful for diagnosing, treating, and/or preventing
PT disorders, such as cancer, infections, cardiovascular and inflammatory
PT diseases.
XX
PS Claim 3; Page 211; 297pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule. The methods
CC and compositions of the present invention are useful for diagnosing,
CC treating, preventing and/or prognosing disorders related to the novel
CC polypeptides, such as neural disorders, immune system disorders (e.g.
CC systemic lupus erythematosus, rheumatoid arthritis, or multiple
CC sclerosis), muscular disorders, respiratory diseases (e.g. nasal
CC vestibulitis, nasal polyps, or sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly or hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure, or end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease or
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis, or
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis or myocardial
CC infarction), endocrine disorders (e.g. Addison's disease or
CC corticosteroid deficiency), reproductive system disorders (e.g.
CC testicular atrophy or dysmenorrhea), gastrointestinal disorders (e.g.
CC dysphagia or irritable bowel syndrome), infectious diseases, and/or
CC cancerous diseases. The polynucleotides can also be used to aid wound
CC healing and epithelial cell proliferation. The present sequence
CC represents a human signal transduction pathway component cDNA

XX Sequence 907 BP; 211 A; 264 C; 255 G; 173 T; 0 U; 4 Other;
SQ
Query Match 1.4%; Score 48; DB 9; Length 907;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3396 TAAATGCGCAAA 3443
Db 859 TAAATGCGCAAA 906
RESULT 3
ABQ54275
ID ABQ54275 standard; cDNA; 1795 BP.
XX AC ABQ54275;
XX
XX 22-AUG-2002 (first entry)
XX
XX Human ovarian antigen HDPSE86 cDNA, SEQ ID NO:155.
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; chromosome 12q23-24;
KW gene; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200200677-A1.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 07-JUN-2001; 2001WO-US018569.
PF
XX
XX 07-JUN-2000; 2000US-0209467P.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Birse CE, Rosen CA;
PI
XX
XX WPI; 2002-147878/19.
DR
XX
XX P-PSDB; ABP41198.
DR
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
PT
XX
XX Claim 1; SEQ ID NO 155; 2922bp; English.
PS
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders,
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1795 BP; 370 A; 548 C; 530 G; 341 T; 0 U; 6 Other;
Query Match 1.4%; Score 48; DB 6; Length 1795;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3396 TAAATGCGCAAA 3443
Db 1747 TAAATGCGCAAA 1794
RESULT 4
ADQ24226
ID ADQ24226 standard; DNA; 2218 BP.
XX AC ADQ24226;
XX
XX 26-AUG-2004 (first entry)
DT
XX
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 7046.
DE
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
KW
XX
XX Homo sapiens.
OS
XX
XX W02004048938-A2.
PN
XX
XX 10-JUN-2004.
PD
XX
XX 26-NOV-2003; 2003WO-US038193.
PF
XX
XX 26-NOV-2002; 2002US-0429739P.
PR
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
PA
XX
XX Aziz N, Ginsburg WM, Zlotnik A;
PI
XX
XX WPI; 2004-441208/41.
DR
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
PT
XX
XX Example 2; SEQ ID NO 7046; 210pp; English.
PS
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual,
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated

CC DNA of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.

XX Sequence 2218 BP; 499 A; 636 C; 646 G; 437 T; 0 U; 0 Other;

SQ Query Match 1.4%; Score 48; DB 12; Length 2218;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGGCAA 3443

DB 1754 TAAATGGCAA 1801

RESULT 5

AAH33312

ID AAH33312 standard; cDNA; 1826 BP.

XX AC AAH33312;

DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen encoding cDNA SEQ ID NO:368.

DE Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 12; ss.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000NO-US026524.

XX 29-SEP-1999; 99US-0157137P.

XX 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

PI WPI; 2001-235357/24.

DR P-PSDB; AAG73881.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 1; Page 2476-2477; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
 CC proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene therapy
 CC and vaccine production. N and P may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate P expression. For
 CC example, N and P may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of P by expressing inactive proteins or to
 CC supplement the patients own production of P. Additionally, N may be used
 CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the proteins. N and P
 CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922

XX Sequence 1826 BP; 593 A; 360 C; 318 G; 554 T; 0 U; 1 Other;

Query Match 1.4%; Score 47; DB 4; Length 1826;

Best Local Similarity 100.0%; Pred. No. 4.1e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3397 AAATGGCAA 3443

DB 1740 AAATGGCAA 1786

RESULT 6

AAF22014

ID AAF22014 standard; DNA; 189 BP.

XX AC AAF22014;

XX 27-MAR-2001 (first entry)

XX Human breast and ovarian cancer associated antigen gene SEQ ID 401.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease; ds.

XX Homo sapiens.

XX WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005881.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

XX P-PSDB; AAB59111.

XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention, treatment
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
 PT neurological diseases.

XX Claim 1; Page 833; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;
 CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
 CC antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
 CC protein sequences are used in the diagnosis of cancer, particularly
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
 CC and agonists may also be used in the diagnosis, prevention and treatment
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus. Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC cardiovascular disorders such as myocardial ischaemias; wound healing;
 CC neurological diseases such as cerebral anoxia and epilepsy; and
 CC infectious diseases

XX Sequence 189 BP; 87 A; 36 C; 20 G; 40 T; 0 U; 6 Other;

Query Match 1.3%; Score 45; DB 3; Length 189;

Best Local Similarity 100.0%; Pred. No. 3.3e-06; Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCAA 3443
Db 104 ATGCAA 148

RESULT 7
ACH24208
ID ACH24208 standard; cDNA; 480 BP.
AC ACH24208;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult ovary cDNA #2588.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 11420; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH2789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030073623
XX
SQ Sequence 480 BP; 129 A; 118 C; 118 G; 98 T; 0 U; 17 Other;
Query Match 1.3%; Score 45; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCAA 3443

Db 208 ATGCAA 252

RESULT 8
AAX51735
ID AAX51735 standard; DNA; 986 BP.
XX
AC AAX51735;
XX
DT 17-JUN-1999 (first entry)
XX
DE DNA encoding a human secreted protein.
XX
KW Human secreted protein; cancer; immune disorder; infection;
KW inflammatory disorder; skin disorder; tumour; atherosclerosis;
KW restenosis; autoimmune disorder; Alzheimer's disease;
KW peripheral neuropathy; trauma; spinal cord injury; allergy;
KW hematopoietic disorder; skeletal disorder; neurological disorder;
KW arthritic disorder; asthma; immunodeficiency disease; AIDS;
KW transplant rejection; ss.
XX
OS Homo sapiens.
XX
PN WO9911293-A1.
XX
PD 11-MAR-1999.
XX
PF 03-SEP-1998; 98WO-US018360.
XX
PR 05-SEP-1997; 97US-0057626P.
PR 05-SEP-1997; 97US-0057663P.
PR 05-SEP-1997; 97US-0057669P.
PR 12-SEP-1997; 97US-0058666P.
PR 12-SEP-1997; 97US-0058667P.
PR 12-SEP-1997; 97US-0058973P.
PR 12-SEP-1997; 97US-0058974P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen GA, Olsen HS;
XX Ebner R, Brewer LA;
XX WPI; 1999-204988/17.
XX P-PSDB; AAY12948.
XX
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. neurological disorders,
PT tumors, immune disorders, inflammation or hematological disorders.
XX
XX Claim 1; Page 171; 215pp; English.
XX
XX AAX51701-55 encode human secreted proteins. The polynucleotides and their
CC corresponding secreted polypeptides are useful for preventing treating
CC or ameliorating medical conditions, e.g. by protein or gene therapy.
CC Pathological conditions can also be diagnosed by determining the amount
CC of the new polypeptides in a sample or by determining the presence of
CC mutations in the new polynucleotides. Specific uses are described for
CC each polynucleotide, based on which tissues they are most highly
CC expressed in, and include developing products for the diagnosis or
CC treatment of cancer, immune disorders, infection, inflammatory disorders,
CC skin disorders, tumors, atherosclerosis, restenosis, autoimmune
CC disorders, Alzheimer's disease, peripheral neuropathies, trauma, spinal
CC cord injuries, allergy, hematopoietic disorders, skeletal disorders,
CC neurological disorders, arthritic disorders, asthma, immunodeficiency
CC diseases, AIDS and transplant rejection. The polypeptides are also useful
CC for identifying their binding partners
XX
SQ Sequence 986 BP; 371 A; 132 C; 177 G; 306 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 2; Length 986;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3399	ATGGCAAA 3443	
Db	883	ATGGCAAA 927	
RESULT 9			
ABQ92587			
ID	ABQ92587	standard; cDNA; 986 BP.	
XX	AC	ABQ92587;	
XX	DT	12-NOV-2002 (first entry)	
XX	DE	Human secreted protein encoding cDNA SEQ ID NO 45.	
XX	KW	Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;	
XX	KW	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;	
XX	KW	antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;	
XX	KW	antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;	
XX	KW	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;	
XX	KW	antiparasitic; cardiac; immune disorder; cardiovascular disorder;	
XX	KW	neurological disease; infection; nephrotropic; gene therapy; vaccine;	
XX	OS	gene; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200257420-A2.	
XX	PD	25-JUL-2002.	
XX	PF	17-JAN-2002; 2002WO-US001109.	
XX	PR	18-JAN-2001; 2001US-0262066P.	
XX	PR	(HUMA-) HUMAN GENOME SCI INC.	
XX	PI	Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen CA, Olsen H;	
XX	PI	Ebner R, Brewer LA;	
XX	XX	WPI; 2002-599716/64.	
XX	XX	P-PSDB; ABP62047.	
XX	XX	New polynucleotides and polypeptides useful for diagnosing, prognosing,	
XX	XX	treating or preventing e.g. neurodegenerative, central nervous system,	
XX	XX	autoimmune, respiratory, reproductive, or inflammatory diseases or	
XX	XX	disorders.	
XX	XX	Claim 1; Page 720; 785pp; English.	
XX	XX	The invention relates to novel genes (ABQ92553-ABQ92607) and proteins	
XX	XX	(ABP62013-ABP62153) useful for preventing, treating or ameliorating	
XX	XX	medical conditions e.g. by protein or gene therapy. The genes are	
XX	XX	isolated from a range of human tissues disclosed in the specification.	
XX	XX	The nucleic acids, proteins, antibodies and (ant)agonists are useful in	
XX	XX	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and	
XX	XX	ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,	
XX	XX	breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune	
XX	XX	disorders e.g. Addison's disease, allergies, autoimmune haemolytic	
XX	XX	anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,	
XX	XX	multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)	
XX	XX	cardiovascular disorders such as myocardial ischaemias; (d) wound healing	
XX	XX	; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)	
XX	XX	infectious diseases such as viral, bacterial, fungal and parasitic	
XX	XX	infections	
XX	XX	Sequence 986 BP; 371 A; 132 C; 177 G; 306 T; 0 U; 0 Other;	
SQ			
Query Match			
Best Local Similarity 1.3%; Score 45; DB 6; Length 986;			
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	3399	ATGGCAAA 3443	
RESULT 10			
ABK33557			
ID	ABK33557	standard; cDNA; 1434 BP.	
XX	AC	ABK33557;	
XX	DT	08-MAY-2002 (first entry)	
XX	DE	cDNA encoding human PRO protein, Seq ID No 43.	
XX	KW	Human; secreted protein; PRO; tumour; lung cancer; colon cancer;	
XX	KW	breast cancer; prostate tumour; rectal tumour; liver tumour;	
XX	KW	pericyte cell proliferation; chondrocyte cell proliferation;	
XX	KW	tumour necrosis factor-alpha; gene; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200208288-A2.	
XX	XX	31-JAN-2002.	
XX	XX	29-JUN-2001; 2001WO-US021066.	
XX	XX	20-JUL-2000; 2000US-0219556P.	
XX	XX	25-JUL-2000; 2000US-0220585P.	
XX	XX	25-JUL-2000; 2000US-0220605P.	
XX	XX	25-JUL-2000; 2000US-0220607P.	
XX	XX	25-JUL-2000; 2000US-0220624P.	
XX	XX	25-JUL-2000; 2000US-0220638P.	
XX	XX	25-JUL-2000; 2000US-0220664P.	
XX	XX	25-JUL-2000; 2000US-0220666P.	
XX	XX	26-JUL-2000; 2000US-0220893P.	
XX	XX	28-JUL-2000; 2000WO-US020710.	
XX	XX	01-AUG-2000; 2000US-0222425P.	
XX	XX	22-AUG-2000; 2000US-0227133P.	
XX	XX	23-AUG-2000; 2000WO-US023522.	
XX	XX	24-AUG-2000; 2000WO-US023328.	
XX	XX	10-NOV-2000; 2000WO-US030873.	
XX	XX	28-NOV-2000; 2000US-0253646P.	
XX	XX	01-DEC-2000; 2000WO-US032678.	
XX	XX	20-DEC-2000; 2000US-00747259.	
XX	XX	20-DEC-2000; 2000WO-US034956.	
XX	XX	28-FEB-2001; 2001WO-US006520.	
XX	XX	01-MAR-2001; 2001WO-US006666.	
XX	XX	22-MAR-2001; 2001US-00816744.	
XX	XX	10-MAY-2001; 2001US-00854208.	
XX	XX	10-MAY-2001; 2001US-00854280.	
XX	XX	25-MAY-2001; 2001WO-US017092.	
XX	XX	(GETH) GENENTECH INC.	
XX	XX	Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;	
XX	XX	Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;	
XX	XX	WPI; 2002-172001/22.	
XX	XX	P-PSDB; AAU83613.	
XX	XX	One hundred and twenty two nucleic acids encoding PRO polypeptides,	
XX	XX	useful for treating a PRO related disorder and for diagnosing tumors such	
XX	XX	as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor	
XX	XX	or liver tumor.	
XX	XX	Claim 2; Fig 43; 359pp; English.	
XX	XX	The invention relates to one hundred and twenty two nucleic acids	
XX	XX	encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides	
XX	XX	encode human secreted proteins. The PRO nucleic acids, polypeptides,	
XX	XX	agonists and antagonists are useful for treating a PRO related disorder.	
XX	XX	The PRO polypeptides are useful for diagnosing tumours, especially lung	

CC cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. The PRO polypeptides are useful for stimulating the
CC proliferation of, or gene expression, in pericyte cells, for stimulating
CC the proliferation or differentiation of chondrocyte cells, for
CC stimulating the release of tumour necrosis factor-alpha from human blood,
CC for stimulating or inhibiting the proliferation of normal human dermal
CC fibroblast cells. The PRO polypeptide may also be used as molecular
CC weight markers and for tissue typing. The PRO nucleic acids have
CC applications in molecular biology, including use as hybridisation probes,
CC and in chromosome and gene mapping. ABK33536-ABK33657 represent human PRO
CC protein coding sequences of the invention

SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 6; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAA 3443

Db 1347 ATGGCAA 1391

RESULT 11

ACA68518
ID ACA68518 standard; cDNA; 1434 BP.

AC ACA68518;

DT 25-JUN-2003 (first entry)

DE Novel human secreted and transmembrane protein PRO9905 cDNA.

KW Human; secreted and transmembrane protein; PRO; cardiant; cytostatic;
KW antiangiogenic; hypotensive; vulnery; antiarteriosclerotic;
KW gene therapy; cardiovascular disorder; endothelial disorder;
KW angiogenic disorder; cardiac hypertrophy; trauma; cancer;
KW age-related macular degeneration; atherosclerosis; hypertension;
KW arterial restenosis; rheumatoid arthritis; angina; myocardial infarction;
KW thrombophlebitis; lymphangitis; tumour angiogenesis; breast carcinoma;
KW liver carcinoma; wound healing; chromosome mapping; gene mapping; gene;
KW ss.

OS Homo sapiens.

PN US2003088063-A1.

XX 08-MAY-2003.

PF 12-AUG-2002; 2002US-00219003.

PR 25-JUL-2000; 2000US-0220664P.

PR 01-JUN-2001; 2001WO-US017800.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;

XX WPI; 2003-393229/37.

DR P-PSDB; ABUS2069.

XX One hundred and eighty seven nucleic acids encoding PRO polypeptides,
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial
PT infarction), endothelial or angiogenic disorders in a mammal.

XX Claim 2; Fig 43; 314pp; English.

XX The invention describes one hundred and eighty seven nucleic acids
CC encoding novel human secreted and transmembrane (PRO) polypeptides. The
CC PRO nucleic acids, polypeptides, agonists and antagonists are useful for

CC treating or diagnosing a cardiovascular, endothelial or angiogenic
CC disorder in a mammal, e.g. cardiac hypertrophy, trauma, cancer, age-
CC related macular degeneration, atherosclerosis, hypertension, arterial
CC restenosis, rheumatoid arthritis, angina, myocardial infarctions,
CC thrombophlebitis, lymphangitis, tumour angiogenesis (such as breast
CC carcinoma and liver carcinoma) and wound healing. The PRO nucleic acids
CC have applications in molecular biology, including use as hybridisation
CC probes, and in chromosome and gene mapping. This sequence encodes a novel
CC human secreted and transmembrane PRO polypeptide

SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 8; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAA 3443

Db 1347 ATGGCAA 1391

RESULT 12

ABT44247

ID ABT44247 standard; cDNA; 1434 BP.

XX ABT44247;

DT 06-NOV-2003 (first entry)

DE Human PRO9905 cDNA.

XX PRO; proliferation; pericyte cell; TNF-alpha; blood; chondrocyte; ss;

KW differentiation; dermal fibroblast; tumour; gene therapy; gene;

KW cytostatic.

OS Homo sapiens.

PN US2003050448-A1.

XX 13-MAR-2003.

XX 28-AUG-2002; 2002US-00230414.

PR 01-JUN-2001; 2001WO-US017800.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;

XX WPI; 2003-521818/49.

DR P-PSDB; ABJ72249.

XX New nucleic acid encoding for a PRO protein, useful for the manufacture
PT of a medicament for diagnosing or treating tumors or for measuring or
PT detecting expression of an associated gene.

XX Claim 2; Fig 43; 315pp; English.

XX The invention relates to a novel isolated nucleic acid encoding a fully
CC defined PRO polypeptide. The molecules of the invention may be useful for
CC stimulating proliferation or gene expression in pericyte cells or the
CC release of TNF-alpha from human blood. Other possible uses include the
CC stimulation or inhibition of chondrocyte proliferation or
CC differentiation, the stimulation of human dermal fibroblast cell
CC proliferation and the detection of the presence of a tumour within a
CC mammal. Furthermore, the nucleic acid may be useful for the manufacture
CC of a medicament for diagnosing or treating a tumour within a mammal or
CC for measuring or detecting the expression of an associated gene, as well
CC as during gene therapy. The current sequence is that of the human PRO
CC cDNA of the invention

```
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|
Db 1347 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391

RESULT 13
ABT44530
ID ABT44530 standard; cDNA; 1434 BP.
XX
AC ABT44530;
XX
XX 06-NOV-2003 (first entry)
XX
DE Human PRO9905 cDNA.
XX
KW PRO; proliferation; gene; pericyte cell; TNF alpha; chondrocyte; blood;
KW tumour necrosis factor; proliferation; differentiation; gene therapy;
KW dermal fibroblast; ss.
XX
OS Homo sapiens.
XX
XX US2003027988-A1.
XX
XX 06-FEB-2003.
XX
XX 26-AUG-2002; 2002US-00227884.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 23-JUN-2001; 2001WO-US021066.
XX
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
XX WPI; 2003-503301/47.
XX
XX P-PSDB; ABJ72377.
XX
XX New PRO protein encoding nucleic acid, useful for preparing PRO
XX polypeptides and anti-PRO antibodies for detecting the presence of a
XX tumor in a mammal.
XX
XX Claim 2; Fig 43; 324pp; English.
XX
XX The invention relates to a novel isolated PRO protein encoding nucleic
XX acid. The nucleic acid of the invention may be useful for preparing PRO
XX polypeptides and anti-PRO antibodies for detecting the presence of a
XX tumour in a mammal. Furthermore, the molecules of the invention may be
XX useful for stimulating proliferation or gene expression in pericyte
XX cells, the release of tumour necrosis factor (TNF)-alpha from human
XX blood, the proliferation or differentiation of chondrocyte cells and for
XX inhibiting the proliferation of normal human dermal fibroblast cells.
XX Finally, the molecules may be utilised during gene therapy. The current
XX sequence is that of the human PRO cDNA of the invention
XX
XX Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|
Db 1347 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391
```

```
RESULT 14
ACD82197
ID ACD82197 standard; cDNA; 1434 BP.
XX
AC ACD82197;
XX
XX 19-SEP-2003 (first entry)
XX
DE Human secreted/transmembrane polypeptide PRO 9905 cDNA.
XX
KW Human; ss; chondrocyte stimulation; TNF-alpha stimulation; gene therapy;
KW human dermal fibroblast stimulation; tumour; tissue typing; gene;
KW affinity purification.
XX
OS Homo sapiens.
XX
XX US2003044934-A1.
XX
XX 06-MAR-2003.
XX
XX 28-AUG-2002; 2002US-00230338.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 29-JUN-2001; 2001WO-US021066.
XX
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
XX WPI; 2003-492274/46.
XX
XX P-PSDB; ABO34272.
XX
XX New transmembrane polypeptides and nucleic acids encoding the
XX polypeptides, useful in gene therapy, in chromosome identification, as
XX chromosome markers, or in generating probes.
XX
XX Claim 2; Fig 43; 315pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding a PRO
XX polypeptide. Nucleic acids that encode PRO can be used to generate either
XX transgenic animals or knock-out animals useful in developing and
XX screening of therapeutically useful reagents. The nucleic acids may also
XX be used in gene therapy for replacing defective gene, in chromosome
XX identification, as chromosome markers, or in generating probes to isolate
XX full length PRO cDNA. The PRO polypeptides are useful for chondrocyte
XX stimulation, TNF-alpha stimulation, human dermal fibroblasts stimulation
XX and for detecting the presence of tumour in an mammal. The PRO
XX polypeptides are useful as molecular markers for protein electrophoresis
XX and the isolated nucleic acids may be used for recombinantly expressing
XX those markers. The PRO polypeptides and nucleic acids may also be used in
XX tissue typing. Anti-PRO antibodies are useful in diagnostic assays for
XX PRO and in affinity purification of PRO from recombinant cell culture or
XX natural sources. The present sequence represents cDNA encoding a human
XX secreted/transmembrane PRO polypeptide
XX
XX Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|
Db 1347 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391

RESULT 15
ABT43903
ID ABT43903 standard; cDNA; 1434 BP.
XX
```

AC ABT43903;
XX
DT 16-OCT-2003 (first entry)
XX
DE Human membrane bound receptor/protein PRO9905 cDNA sequence.
XX
KW Human; PRO; membrane bound protein; membrane bound receptor;
KW cell proliferation; cell migration; cell differentiation;
KW mitogenic factor; survival factor; cytotoxic factor;
KW differentiation factor; neuropeptide; hormone; cell receptor;
KW receptor-ligand interaction; cytostatic; chondrocyte; tumour; gene; ss.
XX
OS Homo sapiens.
XX
FN US2003065147-A1.
XX
PD 03-APR-2003.
XX
PF 29-AUG-2002; 2002US-00232224.
XX
PR 28-JUL-1999; 99US-0146222P.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
PI
XX WPI; 2003-522018/49.
DR P-PSDB; ABJ72079.
XX
XX One hundred and twenty two nucleic acids encoding PRO polypeptides
PT useful for the manufacture of a medicament for diagnosing or treating
PT tumor.
XX
XX Claim 2; Fig 43; 315pp; English.
XX
XX This invention relates to one hundred and twenty two novel nucleic acids
CC encoding human PRO membrane bound proteins or receptors. Extracellular
CC proteins play important roles in the formation, differentiation and
CC maintenance of multicellular organisms. The fate of many individual cells
CC (for example proliferation, migration or differentiation) is typically
CC governed by information received from other cells and the immediate
CC environment. The information is often transmitted by secreted
CC polypeptides (for example mitogenic factors, survival factors, cytotoxic
CC factors, differentiation factors, neuropeptides and hormones) which are
CC received and interpreted by diverse cell receptors or membrane bound
CC proteins. These membrane bound proteins and receptors may be of use as
CC pharmaceutical and diagnostic agents, such as in the blocking of receptor
CC -ligand interactions. The current invention provides the amino acid
CC sequences of novel human membrane bound receptors and proteins, along
CC with the cDNA sequences encoding them. The novel proteins of the
CC invention may have cytostatic activities through the stimulation of
CC chondrocytes. The nucleic acids of the invention may be useful for the
CC manufacture of a medicament for diagnosing or treating a tumour in a
CC mammal. In addition, they may be useful for measuring or detecting the
CC expression of a tumour associated gene. The present sequence is the cDNA
XX sequence encoding a human PRO protein of the invention
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 9; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 3399 ATGCGCAA 3443
Db 1347 ATGCGCAA 1391

RESULT 16
ADB83533
ID ADB83533 standard; cDNA; 1434 BP.
XX
AC ADB83533;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO9905 cDNA.
XX
KW human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnary; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.
XX
XX Homo sapiens.
XX
XX US2003073814-A1.
XX
PD 17-APR-2003.
XX
XX 12-AUG-2002; 2002US-00218849.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 29-JUN-2001; 2001WO-US021066.
XX
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
PI
XX WPI; 2003-644806/61.
DR P-PSDB; ADB83534.
XX
XX New PRO polypeptides and nucleic acids encoding the polypeptides, useful
PT in gene therapy, chromosome identification, tissue typing, or as
PT hybridization probes in chromosome and gene mapping.
XX
XX Claim 2; Fig 43; 315pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1567,
CC PRO1887, PRO1928, PRO4344, PRO1801, PRO4333, PRO3543, PRO4344, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung

CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 9; Length 1434;

Best Local Similarity 100.0%; Pred. No. 2.2e-06;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443

Db 1347 ATGGCAA 1391

RESULT 17

ADB80639

ID ADB80639 standard; cDNA; 1434 BP.

XX ADB80639;

AC ADB80639;

XX 04-DEC-2003 (first entry)

DT 04-DEC-2003 (first entry)

DE Novel human secreted and transmembrane protein PRO9905 cDNA.

XX Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;

XX vulnary; antiarthritic; pericyte cell proliferation;

XX pericyte cell differentiation; chondrocyte cell proliferation;

XX chondrocyte cell differentiation; tumour necrosis factor alpha release;

XX (TNF)-alpha release; dermal fibroblast cell proliferation;

XX dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;

XX colon tumour; breast tumour; prostate tumour; rectal tumour;

XX liver tumour; tissue typing; chromosome mapping; gene mapping;

XX gene therapy.

XX Homo sapiens.

XX US200308068-A1.

XX 08-MAY-2003.

XX 13-AUG-2002; 2002US-00219481.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-657982/62.

XX P-PSDB; ADB80640.

XX One hundred and twenty two nucleic acids encoding PRO polypeptides,

XX useful in gene therapy, chromosome identification, tissue typing, or as

XX hybridization probes in chromosome and gene mapping.

XX Claim 2; Fig 43; 305pp; English.

XX The invention describes an isolated PRO (secreted and transmembrane)

XX polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are

XX useful for stimulating the proliferation of or gene expression in

XX pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful

XX for stimulating the proliferation or differentiation of chondrocyte

XX cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide

CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
CC PRO347, PRO337, PRO536, PRO363, PRO531, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1274, PRO1412,
CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
CC PRO1887, PRO1328, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 9; Length 1434;

Best Local Similarity 100.0%; Pred. No. 2.2e-06;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443

Db 1347 ATGGCAA 1391

RESULT 18

ADB73180

ID ADB73180 standard; cDNA; 1434 BP.

XX ADB73180;

XX 04-DEC-2003 (first entry)

DT 04-DEC-2003 (first entry)

DE Novel human secreted and transmembrane protein PRO9905 cDNA.

XX human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;

XX vulnary; antiarthritic; pericyte cell proliferation;

XX pericyte cell differentiation; chondrocyte cell proliferation;

XX chondrocyte cell differentiation; tumour necrosis factor alpha release;

XX (TNF)-alpha release; dermal fibroblast cell proliferation;

XX dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;

XX colon tumour; breast tumour; prostate tumour; rectal tumour;

XX liver tumour; tissue typing; chromosome mapping; gene mapping;

XX gene therapy.

XX Homo sapiens.

XX US200309668-A1.

XX 22-MAY-2003.

XX 29-AUG-2002; 2002US-00232223.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

(GETH) GENENTECH INC.

Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PU;
 Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 WPI; 2003-765525/72.
 P-PSDB; ADB73181.

New isolated PRO polypeptides useful as molecular weight markers in
 protein electrophoresis, useful for tissue typing, and for treating
 arthritis and tumors.

Claim 2; Fig 43; 308pp; English.

The invention describes an isolated PRO (secreted and transmembrane)
 polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
 useful for stimulating the proliferation of or gene expression in
 pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
 for stimulating the proliferation or differentiation of chondrocyte
 cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
 are useful for stimulating the release of tumour necrosis factor (TNF)-
 alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
 PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
 PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
 PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
 PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
 PRO1343, PRO1376, PRO1387, PRO1349, PRO1474, PRO1917, PRO1760, PRO1567,
 PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
 PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
 stimulating the proliferation of normal human dermal fibroblasts cells.
 PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
 PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
 inhibiting the proliferation of normal human dermal fibroblast cells. PRO
 polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
 are useful for detecting the presence of tumour in a mammal which
 involves comparing the level of expression of the above PRO polypeptides
 in a test sample of cells taken from the mammal, and a control sample of
 normal cells of the same cell type, where a higher level of expression of
 the PRO polypeptides in the test sample as compared to the control sample
 is indicative of the presence of tumour in the mammal. The tumour is lung
 tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 liver tumour. (I) is useful as molecular weight markers, for tissue
 typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
 useful for chromosome and gene mapping or gene therapy. (II) is useful
 for generating transgenic animals or knock-out animals which are useful
 for screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
 Query Match 1.3%; Score 45; DB 9; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
 |||||||
 Db 1347 ATGCGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391
 |||||||

RESULT 19
 ADB78262
 ID ADB78262 standard; cDNA; 1434 BP.

XX
 AC ADB78262;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO9905 cDNA.

XX
 KW Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnary; antiarthritic; pericyte cell proliferation;

KW pericyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX Homo sapiens.
 OS
 XX
 XX US2003092889-A1.
 PN
 XX
 XX 15-MAY-2003.
 PD
 XX
 XX 13-AUG-2002; 2002US-00219478.
 PF
 XX
 XX 01-JUN-2001; 2001WO-US017800.
 PR
 XX 29-JUN-2001; 2001WO-US021066.
 PR
 XX 09-APR-2002; 2002US-00119480.
 PR
 XX (GETH) GENENTECH INC.
 PA
 XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PU;
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 PI WPI; 2003-765495/72.
 DR P-PSDB; ADB78263.

New isolated PRO polypeptide useful for tissue typing, gene therapy, as
 molecular weight markers in protein electrophoresis, and for treating
 arthritis and tumors.

Claim 2; Fig 43; 308pp; English.

The invention describes an isolated PRO (secreted and transmembrane)
 polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
 useful for stimulating the proliferation of or gene expression in
 pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
 for stimulating the proliferation or differentiation of chondrocyte
 cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
 are useful for stimulating the release of tumour necrosis factor (TNF)-
 alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
 PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
 PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
 PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
 PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
 PRO1343, PRO1376, PRO1387, PRO1349, PRO1474, PRO1917, PRO1760, PRO1567,
 PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
 PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
 stimulating the proliferation of normal human dermal fibroblasts cells.
 PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
 PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
 inhibiting the proliferation of normal human dermal fibroblast cells. PRO
 polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
 are useful for detecting the presence of tumour in a mammal which
 involves comparing the level of expression of the above PRO polypeptides
 in a test sample of cells taken from the mammal, and a control sample of
 normal cells of the same cell type, where a higher level of expression of
 the PRO polypeptides in the test sample as compared to the control sample
 is indicative of the presence of tumour in the mammal. The tumour is lung
 tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 liver tumour. (I) is useful as molecular weight markers, for tissue
 typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
 useful for chromosome and gene mapping or gene therapy. (II) is useful
 for generating transgenic animals or knock-out animals which are useful
 for screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
 Query Match 1.3%; Score 45; DB 9; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
 |||||||
 Db 1347 ATGCGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391
 |||||||

RESULT 19
 ADB78262
 ID ADB78262 standard; cDNA; 1434 BP.

XX
 AC ADB78262;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO9905 cDNA.

XX
 KW Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnary; antiarthritic; pericyte cell proliferation;

Best Local Similarity 100.0%; Pred. No. 2.2e-06; Mismatches 0; Indels 0; Gaps 0;

3399 ATGGCAAA 3443
1347 ATGGCAAA 1391

Query Match 1.3%; Score 45; DB 10; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAA 3443
DB 1347 ATGGCAAA 1391

RESULT 20
ADB84910
ID ADB84910 standard; cDNA; 1434 BP.
AC ADB84910;
DT 04-DEC-2003 (first entry)
DE Human PRO polynucleotide #22.
XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
KW tumour; cancer; lung; colon; breast; prostate; rectum; liver;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
KW pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
KW arthritis; sports injury; cytostatic; antiarthritic.
OS Homo sapiens.
XX US2003073817-A1.
PN 17-APR-2003.
PD 26-AUG-2002; 2002US-00227883.
XX 01-JUN-2000; 2000US-022425P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX (GETH) GENENTECH INC.
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI: 2003-730024/69.
DR P-PSDB; ADB84911.
XX New PRO polypeptides and nucleic acids encoding the polypeptides, useful
PT e.g. in gene therapy, disease diagnosis, chromosome identification and
PT tissue typing.
XX Claim 2; Fig 43; 314pp; English.
XX The invention relates to human PRO polypeptides (secreted and
XX transmembrane polypeptides) and the PRO polynucleotides encoding them.
XX The PRO polypeptides and polynucleotides are useful as pharmaceuticals,
XX diagnostics, biosensors or bioreactors. They are particularly useful for
XX detecting tumours (e.g. lung tumour, colon tumour, breast tumour,
XX prostate tumour, rectal tumour or liver tumour) in a mammal, for
XX stimulating the release of tumour necrosis factor (TNF)-alpha from human
XX blood, for stimulating the proliferation or differentiation of
XX chondrocyte cells, for stimulating the proliferation of or gene
XX expression in pericyte cells or for stimulating the proliferation of
XX normal human dermal fibroblasts. The PRO nucleic acids are useful as
XX hybridisation probes, in chromosome and gene mapping, in generating
XX antisense RNA and DNA, in preparing PRO polypeptides by recombinant
XX technology, in generating transgenic animals or knock-out animals which
XX may be used in the development and screening of therapeutically useful
XX reagents, in gene therapy, in chromosome identification, as chromosome
XX markers and in generating probes. The PRO polypeptides, or anti-PRO
XX antibodies, are useful for preparing a medicament for treating a
XX condition which is responsive to the PRO polypeptides or anti-PRO
XX antibodies, such as pericyte-associated tumours and bone and/or cartilage
XX disorders (e.g. arthritis, sports injuries), involving inducing the re-
XX differentiation of chondrocytes. The PRO polypeptides are useful as

molecular markers for protein electrophoresis, and in tissue typing. This
sequence represents a human PRO polynucleotide of the invention.

Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 10; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAAA 3443
DB 1347 ATGGCAAA 1391

RESULT 21
ADB78016
ID ADB78016 standard; cDNA; 1434 BP.
XX ADB78016;
AC ADB78016;
DT 04-DEC-2003 (first entry)
XX Novel human secreted and transmembrane protein PRO9905 cDNA.
DE Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnery; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.
XX Homo sapiens.
XX US2003092886-A1.
PN 15-MAY-2003.
PD 09-AUG-2002; 2002US-00216165.
XX 25-JUL-2000; 2000US-0220607P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX (GETH) GENENTECH INC.
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI: 2003-765494/72.
DR P-PSDB; ADB78017.
XX Novel isolated PRO polypeptide useful for tissue typing. Gene therapy, as
PT molecular weight markers in protein electrophoresis, for treating
PT arthritis, tumor.
XX Claim 2; Fig 43; 308pp; English.
XX The invention describes an isolated PRO (secreted and transmembrane)
XX polypeptide (1). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
XX useful for stimulating the proliferation of or gene expression in
XX pericyte cells. PRO3357, PRO229, PRO1272 or PRO4405 polypeptide are useful
XX for stimulating the proliferation or differentiation of chondrocyte
XX cells. PRO331, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
XX are useful for stimulating the release of tumour necrosis factor (TNF)-
XX alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO214,
XX PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO1080,
XX PRO1478, PRO1134, PRO826, PRO809, PRO1071, PRO1411, PRO1309,
XX PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1412,
XX PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,

CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;

Best Local Similarity 100.0%; Pred. No. 2.2e-06;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCGAA 3443

Db 1347 ATGCGAA 1391

RESULT 22

ADB87082

ID ADB87082 standard; cDNA; 1434 BP.

XX

AC ADB87082;

XX 04-DEC-2003 (first entry)

XX Human PRO polynucleotide #22.

XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
XX tumour; cancer; lung; colon; breast; prostate; rectum; liver;
XX tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
XX pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
XX arthritis; sports injury; cytostatic; antiarthritic.

XX Homo sapiens.

XX US2003088067-A1.

XX 08-MAY-2003.

XX 13-AUG-2002; 2002US-00219479.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PI;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-657981/62.

XX P-PSDB; ADB87083.

XX One hundred and twenty two nucleic acids encoding PRO polypeptides, or as

XX useful in gene therapy, chromosome identification, tissue typing, or as

PT hybridization probes in chromosome and gene mapping.

XX Claim 2; Fig 43; 314pp; English.

PS The invention relates to human PRO polypeptides (secreted and

XX transmembrane polypeptides) and the PRO polynucleotides encoding them.

CC The PRO polypeptides and polynucleotides are useful as pharmaceuticals,

CC diagnostics, biosensors or bioreactors. They are particularly useful for

CC detecting tumours (e.g. lung tumour, colon tumour, breast tumour,

CC prostate tumour, rectal tumour or liver tumour) in a mammal, for

CC stimulating the release of tumour necrosis factor (TNF)-alpha from human

CC blood, for stimulating the proliferation or differentiation of

CC chondrocyte cells, for stimulating the proliferation of or gene

CC expression in pericyte cells or for stimulating the proliferation of

CC normal human dermal fibroblasts. The PRO nucleic acids are useful as

CC hybridisation probes, in chromosome and gene mapping, in generating

CC antisense RNA and DNA, in preparing PRO polypeptides by recombinant

CC technology, in generating transgenic animals or knock-out animals which

CC may be used in the development and screening of therapeutically useful

CC reagents, in gene therapy, in chromosome identification, as chromosome

CC markers and in generating probes. The PRO polypeptides, or anti-PRO

CC antibodies, are useful for preparing a medicament for treating a

CC condition which is responsive to the PRO polypeptides or anti-PRO

CC antibodies, such as pericyte-associated tumours and bone and/or cartilage

CC disorders (e.g. arthritis, sports injuries), involving inducing the re-

CC differentiation of chondrocytes. The PRO polypeptides are useful as

CC molecular markers for protein electrophoresis, and in tissue typing. This

CC sequence represents a human PRO polynucleotide of the invention.

XX

SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;

Best Local Similarity 100.0%; Pred. No. 2.2e-06;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGCGAA 3443

Db 1347 ATGCGAA 1391

RESULT 23

ADB84664

ID ADB84664 standard; cDNA; 1434 BP.

XX

AC ADB84664;

XX 04-DEC-2003 (first entry)

XX Human PRO polynucleotide #22.

XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;

XX tumour; cancer; lung; colon; breast; prostate; rectum; liver;

XX tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;

XX pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;

XX arthritis; sports injury; cytostatic; antiarthritic.

XX Homo sapiens.

XX US2003092890-A1.

XX 15-MAY-2003.

XX 14-AUG-2002; 2002US-00219536.

XX 28-JUL-1999; 99US-0146222P.

XX 24-FEB-2000; 2000WO-US005004.

XX 02-MAR-2000; 2000WO-US005841.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PI;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-657981/62.

XX P-PSDB; ADB87083.

XX One hundred and twenty two nucleic acids encoding PRO polypeptides, or as

XX useful in gene therapy, chromosome identification, tissue typing, or as

AC	ABD72934;
XX	04-DEC-2003 (first entry)
XX	Novel human secreted and transmembrane protein PRO9905 cDNA.
DE	human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
XX	vulnary; antiarthritic; pericyte cell proliferation;
KW	pericyte cell differentiation; chondrocyte cell proliferation;
KW	chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW	(TNF)-alpha release; dermal fibroblast cell proliferation;
KW	dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW	colon tumour; breast tumour; prostate tumour; rectal tumour;
KW	liver tumour; tissue typing; chromosome mapping; gene mapping;
KW	gene therapy.
OS	Homo sapiens.
XX	
XX	US2003092887-A1.
PN	
XX	15-MAY-2003.
PD	
XX	
XX	12-AUG-2002; 2002US-00218956.
PF	
XX	
XX	29-JUN-2001; 2001WO-US021066.
PR	
XX	09-APR-2002; 2002US-00119480.
PR	
XX	(GETH) GENENTECH INC.
PA	
XX	
XX	Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI	Grimaldi JC, Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI;
PI	
XX	WPI; 2003-777258/73.
DR	
DR	P-PSDB; ADB72935.
XX	
XX	Novel isolated PRO polypeptide useful for tissue typing, gene therapy, as
PT	molecular weight markers, for treating arthritis, tumor.
PT	
XX	Claim 2; Fig 43; 308pp; English.
PS	
XX	
CC	The invention describes an isolated PRO (secreted and transmembrane)
CC	polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC	useful for stimulating the proliferation of or gene expression in
CC	pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC	for stimulating the proliferation or differentiation of chondrocyte
CC	cells. PRO331, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
CC	are useful for stimulating the release of tumour necrosis factor (TNF)-
CC	alpha from human blood. PRO982, PRO357, PRO1306, PRO1419, PRO214,
CC	PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
CC	PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC	PRO1025, PRO1181, PRO1136, PRO1186, PRO1192, PRO1244, PRO1274, PRO1412,
CC	PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
CC	PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
CC	PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
CC	PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC	stimulating the proliferation of normal human dermal fibroblasts cells.
CC	PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC	PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC	inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC	polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC	are useful for detecting the presence of tumour in a mammal which
CC	involves comparing the level of expression of the above PRO polypeptides
CC	in a test sample of cells taken from the mammal, and a control sample of
CC	normal cells of the same cell type, where a higher level of expression of
CC	the PRO polypeptides in the test sample as compared to the control sample
CC	is indicative of the presence of tumour in the mammal. The tumour is lung
CC	tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC	liver tumour. (I) is useful as molecular weight markers, for tissue
CC	typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC	useful for chromosome and gene mapping or gene therapy. (II) is useful
CC	for generating transgenic animals or knock-out animals which are useful
CC	screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC	is useful for treating bone and/or cartilage disorders (e.g. arthritis,
CC	

CC antibodies, are useful for preparing a medicament for treating a
 CC condition which is responsive to the PRO polypeptides or anti-PRO
 CC antibodies, such as pericyte-associated tumours and bone and/or cartilage
 CC disorders (e.g. arthritis, sports injuries), involving inducing the re-
 CC differentiation of chondrocytes. The PRO polypeptides are useful as
 CC molecular markers for protein electrophoresis, and in tissue typing. This
 CC sequence represents a human PRO polynucleotide of the invention.

XX
 SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
 |||||
 Db 1347 ATGGCAA 1391

RESULT 27

ADC21762

ID ADC21762 standard; cDNA; 1434 BP.

XX AC ADC21762;

XX DT 18-DEC-2003 (first entry)

XX DE Human PRO polynucleotide #22.

XX KW Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
 KW tumour; cancer; lung; colon; breast; prostate; rectum; liver;
 KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
 KW pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
 KW arthritis; sports injury; cytostatic; antiarthritic.

XX OS Homo sapiens.

XX PN US2003096969-A1.

XX PD 22-MAY-2003.

XX PF 29-AUG-2002; 2002US-00232225.

XX PR 02-JUN-2000; 2000WO-US015264.

XX PR 05-JUN-2000; 2000US-0209832P.

XX PR 20-JUN-2000; 2000US-0212901P.

XX PR 22-JUN-2000; 2000US-0213807P.

XX PR 20-JUL-2000; 2000US-0219356P.

XX PR 25-JUL-2000; 2000US-0220585P.

XX PR 25-JUL-2000; 2000US-0220605P.

XX PR 25-JUL-2000; 2000US-0220607P.

XX PR 25-JUL-2000; 2000US-0220624P.

XX PR 25-JUL-2000; 2000US-0220638P.

XX PR 25-JUL-2000; 2000US-0220664P.

XX PR 25-JUL-2000; 2000US-0220666P.

XX PR 26-JUL-2000; 2000US-0220893P.

XX PR 01-AUG-2000; 2000US-0222425P.

XX PR 23-AUG-2000; 2000US-0227133P.

XX PR 24-AUG-2000; 2000WO-US023328.

XX PR 10-NOV-2000; 2000WO-US030873.

XX PR 28-NOV-2000; 2000US-0253646P.

XX PR 01-DEC-2000; 2000WO-US032678.

XX PR 20-DEC-2000; 2000US-00747259.

XX PR 28-DEC-2000; 2000WO-US034956.

XX PR 28-FEB-2001; 2001WO-US006520.

XX PR 25-MAY-2001; 2001WO-US017092.

XX PR 01-JUN-2001; 2001WO-US017800.

XX PR 29-JUN-2001; 2001WO-US021066.

XX PR 09-APR-2002; 2002US-00119480.

XX PA (GETH) GENENTECH INC.

XX

PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 XX WPI; 2003-765526/72.
 DR P-PSDB; ADC21763.
 XX
 PT Novel isolated PRO polypeptide useful for tissue typing, as molecular
 PT weight markers in protein electrophoresis, for treating arthritis, tumor.
 XX
 XX Claim 2; Fig 43; 308pp; English.

XX The invention relates to human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the PRO polynucleotides encoding them.
 CC The PRO polypeptides and polynucleotides are useful as pharmaceuticals,
 CC diagnostics, biosensors or bioeffectors. They are particularly useful for
 CC detecting tumours (e.g. lung tumour, colon tumour, breast tumour,
 CC prostate tumour, rectal tumour or liver tumour) in a mammal, for
 CC stimulating the release of tumour necrosis factor (TNF)-alpha from human
 CC blood, for stimulating the proliferation or differentiation of
 CC chondrocyte cells, for stimulating the proliferation of or gene
 CC expression in pericyte cells or for stimulating the proliferation of
 CC normal human dermal fibroblasts. The PRO nucleic acids are useful as
 CC hybridisation probes, in chromosome and gene mapping, in generating
 CC antisense RNA and DNA, in preparing PRO polypeptides by recombinant
 CC technology, in generating transgenic animals or knock-out animals which
 CC may be used in the development and screening of therapeutically useful
 CC reagents, in gene therapy, in chromosome identification, as chromosome
 CC markers and in generating probes. The PRO polypeptides, or anti-PRO
 CC antibodies, are useful for preparing a medicament for treating a
 CC condition which is responsive to the PRO polypeptides or anti-PRO
 CC antibodies, such as pericyte-associated tumours and bone and/or cartilage
 CC disorders (e.g. arthritis, sports injuries), involving inducing the re-
 CC differentiation of chondrocytes. The PRO polypeptides are useful as
 CC molecular markers for protein electrophoresis, and in tissue typing. This
 CC sequence represents a human PRO polynucleotide of the invention.

XX SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
 |||||
 Db 1347 ATGGCAA 1391

RESULT 28

ADC49793

ID ADC49793 standard; cDNA; 1434 BP.

XX AC ADC49793;

XX DT 18-DEC-2003 (first entry)

XX DE Novel human secreted and transmembrane protein PRO9905 cDNA.

XX KW human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnery; antiarthritic; pericyte cell proliferation;
 KW pericyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX OS Homo sapiens.

XX PN US2003088064-A1.

XX PD 08-MAY-2003.

XX

PF 14-AUG-2002; 2002US-00219075.
XX
PR 25-JUL-2000; 2000US-0220605P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
PA (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WJ;
XX WPI; 2003-801154/75.
DR P-PSDB; ADC49794.
DR
XX
PT New secreted and transmembrane PRO polypeptide useful for preparing a
PT medicament for treating a condition that is responsive to the PRO
PT polypeptide or anti-PRO antibody, e.g. cancer.
XX
PS Claim 2; SEQ ID NO 43; 314pp; English.
XX
CC The invention describes an isolated PRO (secreted and transmembrane)
CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
CC useful for stimulating the proliferation of or gene expression in
CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
CC are useful for stimulating the release of tumour necrosis factor (TNF)-
CC alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
CC PRO247, PRO337, PRO526, PRO363, PRO831, PRO1083, PRO840, PRO1080,
CC PRO1478, PRO1134, PRO1126, PRO1186, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1567,
CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
CC stimulating the proliferation of normal human dermal fibroblasts cells.
CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
CC are useful for detecting the presence of tumour in a mammal which
CC involves comparing the level of expression of the above PRO polypeptides
CC in a test sample of cells taken from the mammal, and a control sample of
CC normal cells of the same cell type, where a higher level of expression of
CC the PRO polypeptides in the test sample as compared to the control sample
CC is indicative of the presence of tumour in the mammal. The tumour is lung
CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
CC liver tumour. (I) is useful as molecular weight markers, for tissue
CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;
Query Match 1.3%; Score 45; DB 10; Length 1434;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3399 ATGCCAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||||
Db 1347 ATGCCAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1391
|||||
RESULT 29
ADC48992
ID ADC48992 standard; cDNA; 1434 BP.
XX
AC ADC48992;

XX 18-DEC-2003 (first entry)
XX Novel human secreted and transmembrane protein PRO9905 CDNA.
XX
XX human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
XX vulnary; antiarthritic; pericyte cell proliferation;
XX pericyte cell differentiation; chondrocyte cell proliferation;
XX chondrocyte cell differentiation; tumour necrosis factor alpha release;
XX (TNF)-alpha release; dermal fibroblast cell proliferation;
XX dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
XX colon tumour; breast tumour; prostate tumour; rectal tumour;
XX liver tumour; tissue typing; chromosome mapping; gene mapping;
XX gene therapy.
XX
XX Homo sapiens.
XX
XX US2003088070-A1.
XX
XX 08-MAY-2003.
XX
XX 28-AUG-2002; 2002US-00230260.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 29-JUN-2001; 2001WO-US021066.
XX
XX 09-APR-2002; 2002US-00119480.
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WJ;
XX WPI; 2003-801155/75.
XX P-PSDB; ADC48993.
XX
XX New PRO polypeptides and nucleic acids encoding the polypeptides, useful
XX in gene therapy, chromosome identification, tissue typing, or as
XX hybridization probes in chromosome and gene mapping.
XX
XX Claim 2; SEQ ID NO 43; 315pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
XX polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
XX useful for stimulating the proliferation of or gene expression in
XX pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
XX for stimulating the proliferation or differentiation of chondrocyte
XX cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
XX are useful for stimulating the release of tumour necrosis factor (TNF)-
XX alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
XX PRO247, PRO337, PRO526, PRO363, PRO831, PRO1083, PRO840, PRO1080,
XX PRO1478, PRO1134, PRO1126, PRO1186, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
XX PRO1025, PRO1181, PRO1126, PRO1186, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
XX PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
XX PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1567,
XX PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
XX PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
XX stimulating the proliferation of normal human dermal fibroblasts cells.
XX PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
XX PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
XX inhibiting the proliferation of normal human dermal fibroblast cells. PRO
XX polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
XX are useful for detecting the presence of tumour in a mammal which
XX involves comparing the level of expression of the above PRO polypeptides
XX in a test sample of cells taken from the mammal, and a control sample of
XX normal cells of the same cell type, where a higher level of expression of
XX the PRO polypeptides in the test sample as compared to the control sample
XX is indicative of the presence of tumour in the mammal. The tumour is lung
XX tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
XX liver tumour. (I) is useful as molecular weight markers, for tissue
XX typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
XX useful for chromosome and gene mapping or gene therapy. (II) is useful
XX for generating transgenic animals or knock-out animals which are useful
XX screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
XX is useful for treating bone and/or cartilage disorders (e.g., arthritis,
XX sport injuries). This sequence encodes a human secreted and transmembrane
XX PRO polypeptide.

CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

XX
 SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
 |||||
 Db 1347 ATGGCAA 1391

RESULT 30

ADC49509

ID ADC49509 standard; cDNA; 1434 BP.

AC ADC49509;

XX 18-DEC-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO9905 cDNA.

XX human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnery; antiarthritic; pericyte cell proliferation;
 KW pericyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX Homo sapiens.

XX US2003088071-A1.

XX 08-MAY-2003.

XX 29-AUG-2002; 2002US-00232231.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-801156/75.

DR P-PSDB; ADC49510.

XX New PRO polypeptides and nucleic acids encoding the polypeptides, useful
 PT in gene therapy, chromosome identification, tissue typing, or as
 PT hybridization probes in chromosome and gene mapping.

XX Claim 2; SEQ ID NO 43; 315pp; English.

XX The invention describes an isolated PRO (secreted and transmembrane)
 CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
 CC useful for stimulating the proliferation of or gene expression in
 CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
 CC for stimulating the proliferation or differentiation of chondrocyte
 CC cells. PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
 CC are useful for stimulating the release of tumour necrosis factor (TNF)-
 CC alpha from human blood. PRO982, PRO357, PRO1306, PRO1419, PRO214,
 CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
 CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
 CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
 CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,

CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
 CC PRO1887, PRO1928, PRO4341, PRO1801, PRO4333, PRO3543, PRO3444, PRO4322,
 CC PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
 CC stimulating the proliferation of normal human dermal fibroblasts cells.
 CC PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
 CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
 CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
 CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
 CC are useful for detecting the presence of expression of the above PRO polypeptides
 CC involves comparing the level of expression of the above PRO polypeptides
 CC in a test sample of cells taken from the mammal, and a control sample of
 CC normal cells of the same cell type, where a higher level of expression of
 CC the PRO polypeptides in the test sample as compared to the control sample
 CC is indicative of the presence of tumour in the mammal. The tumour is lung
 CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. (I) is useful as molecular weight markers, for tissue
 CC typing or as therapeutic agents. A polynucleotide (II) encoding (I) is
 CC useful for chromosome and gene mapping or gene therapy. (II) is useful
 CC for generating transgenic animals or knock-out animals which are useful
 CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

XX SQ Sequence 1434 BP; 320 A; 441 C; 391 G; 282 T; 0 U; 0 Other;

Query Match 1.3%; Score 45; DB 10; Length 1434;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443

Db 1347 ATGGCAA 1391

Search completed: March 12, 2005, 03:12:39

Job time : 1175 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2005, 02:31:53 ; Search time 388 Seconds
(without alignments)
14519.847 Million cell updates/sec

Title: US-10-627-132-29

Perfect score: 3443

Sequence: 1 ctgcgtgcctgcctgcga.....aaaaaaaaaaaaaaaaaaaa 3443

Scoring table:

OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 25

Total number of hits satisfying chosen parameters: 13194

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	1.3	1445	3	US-09-697-367-19
2	44	1.3	1445	4	US-09-918-909A-19
3	44	1.3	1530	4	US-09-811-361-19
4	44	1.3	1560	4	US-09-500-495A-5
5	44	1.3	1653	3	US-09-345-469-2
6	44	1.3	1720	3	US-09-227-357-139
7	44	1.3	2311	4	US-09-614-912-91
8	44	1.3	2806	3	US-09-653-839-9
9	44	1.3	2806	4	US-10-202-619-9
10	44	1.3	2852	3	US-09-027-137-2
11	44	1.3	2852	3	US-09-344-441-2
12	44	1.3	3746	4	US-10-160-719A-29
13	44	1.3	3746	4	US-10-160-719A-49
14	44	1.3	3773	4	US-10-160-719A-5
15	44	1.3	3969	4	US-10-160-719A-17
16	44	1.3	3969	4	US-10-160-719A-37
17	44	1.3	5962	6	US-09-625-5
18	44	1.3	5962	6	US-09-625-5
19	44	1.3	5975	1	US-08-404-354B-1
20	44	1.3	5975	1	US-08-314-083B-1
21	44	1.3	5975	1	US-08-435-675B-1
22	44	1.3	5975	1	US-08-336-257A-3
23	44	1.3	5975	3	US-08-884-599-1
24	43	1.2	140	4	US-09-621-976-16833
25	43	1.2	141	4	US-09-621-976-16834
26	43	1.2	194	4	US-09-621-976-801
27	43	1.2	286	4	US-09-621-976-17670

28	43	1.2	308	2	US-08-721-488-4	Sequence 4, Appli
29	43	1.2	674	4	US-09-620-405B-465	Sequence 465, App
30	43	1.2	674	4	US-09-433-826B-465	Sequence 465, App
31	43	1.2	674	4	US-09-604-287A-465	Sequence 465, App
32	43	1.2	674	4	US-09-834-759-465	Sequence 465, App
33	43	1.2	674	4	US-09-590-751A-465	Sequence 465, App
34	43	1.2	674	4	US-09-551-621-465	Sequence 465, App
35	43	1.2	1134	3	US-09-248-335-29	Sequence 2, Appli
36	43	1.2	1196	4	US-09-065-040-2	Sequence 110, App
37	43	1.2	1525	4	US-09-461-325-110	Sequence 110, App
38	43	1.2	1525	4	US-10-012-542-110	Sequence 110, App
39	43	1.2	1525	4	US-10-115-123-110	Sequence 110, App
40	43	1.2	1559	4	US-09-489-847-42	Sequence 42, Appl
41	43	1.2	1949	4	US-09-461-325-26	Sequence 26, Appl
42	43	1.2	1949	4	US-10-012-542-26	Sequence 26, Appl
43	43	1.2	1949	4	US-10-115-123-26	Sequence 26, Appl
44	43	1.2	2209	1	US-08-514-014-1	Sequence 1, Appli
45	43	1.2	2209	2	US-08-833-823-1	Sequence 1, Appli
46	43	1.2	2280	3	US-08-813-150-1	Sequence 1, Appli
47	43	1.2	2280	4	US-09-546-553-1	Sequence 1, Appli
48	43	1.2	2287	4	US-09-949-016-134	Sequence 134, App
49	43	1.2	2442	4	US-09-575-081B-3	Sequence 3, Appli
50	43	1.2	2950	4	US-09-489-847-92	Sequence 92, Appl
51	43	1.2	2964	4	US-09-578-063-25	Sequence 25, Appl
52	43	1.2	4419	4	US-09-620-312B-187	Sequence 187, App
53	43	1.2	7125	4	US-09-949-016-12359	Sequence 12359, A
54	43	1.2	7156	4	US-09-949-016-16710	Sequence 16710, A
55	42	1.2	133	4	US-09-621-976-8845	Sequence 8845, Ap
56	42	1.2	138	4	US-09-621-976-8112	Sequence 8112, Ap
57	42	1.2	139	4	US-09-621-976-8632	Sequence 8632, Ap
58	42	1.2	150	4	US-09-621-976-8656	Sequence 8656, Ap
59	42	1.2	165	4	US-09-621-976-8127	Sequence 8127, Ap
60	42	1.2	176	4	US-09-621-976-13903	Sequence 13903, A
61	42	1.2	191	4	US-09-621-976-18332	Sequence 18332, A
62	42	1.2	225	4	US-09-621-976-18371	Sequence 18371, A
63	42	1.2	233	4	US-09-621-976-10675	Sequence 10675, A
64	42	1.2	241	4	US-09-621-976-16340	Sequence 16340, A
65	42	1.2	249	4	US-09-621-976-19144	Sequence 19144, A
66	42	1.2	269	4	US-09-621-976-17610	Sequence 17610, A
67	42	1.2	362	4	US-09-621-976-10457	Sequence 10457, A
68	42	1.2	491	4	US-09-311-021-191	Sequence 191, App
69	42	1.2	515	3	US-08-589-028-1	Sequence 1, Appli
70	42	1.2	515	3	US-08-785-271-1	Sequence 1, Appli
71	42	1.2	515	3	US-08-785-271-1	Sequence 1, Appli
72	42	1.2	630	1	US-08-185-414E-1	Sequence 13081, A
73	42	1.2	708	4	US-09-707-767-13081	Sequence 13081, A
74	42	1.2	787	4	US-09-621-976-1878	Sequence 1878, Ap
75	42	1.2	806	3	US-08-955-629C-1	Sequence 1, Appli
76	42	1.2	831	4	US-09-904-615-25	Sequence 25, Appl
77	42	1.2	857	1	US-08-308-883-1	Sequence 1, Appli
78	42	1.2	857	1	US-08-730-163-1	Sequence 1, Appli
79	42	1.2	857	3	US-08-256-799-1	Sequence 1, Appli
80	42	1.2	857	3	US-08-462-437-1	Sequence 1, Appli
81	42	1.2	941	4	US-09-205-258-186	Sequence 186, App
82	42	1.2	971	4	US-09-800-729-49	Sequence 49, Appl
83	42	1.2	985	4	US-09-322-409-25	Sequence 25, Appl
84	42	1.2	985	4	US-09-322-409-27	Sequence 27, Appl
85	42	1.2	985	4	US-09-451-527-25	Sequence 25, Appl
86	42	1.2	985	4	US-09-451-527-27	Sequence 27, Appl
87	42	1.2	1023	1	US-08-252-266B-16	Sequence 16, Appl
88	42	1.2	1036	4	US-09-205-258-86	Sequence 86, Appl
89	42	1.2	1050	4	US-09-482-273-58	Sequence 58, Appl
90	42	1.2	1066	1	US-08-157-101A-4	Sequence 4, Appli
91	42	1.2	1181	3	US-09-149-476-310	Sequence 310, App
92	42	1.2	1212	3	US-09-149-476-186	Sequence 186, App
93	42	1.2	1307	2	US-08-960-022-17	Sequence 17, Appl
94	42	1.2	1358	4	US-09-949-016-463	Sequence 463, App
95	42	1.2	1359	3	US-09-387-574-11	Sequence 11, Appl
96	42	1.2	1359	3	US-09-668-096-11	Sequence 11, Appl
97	42	1.2	1405	4	US-09-668-097A-15	Sequence 15, Appl
98	42	1.2	1459	4	US-09-537-654-3	Sequence 3, Appli
99	42	1.2	1503	3	US-08-323-454A-5	Sequence 5, Appli
100	42	1.2	1534	1	US-08-300-903A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-697-367-19
; Sequence 19, Application US/09697367
; Patent No. 6323015
; GENERAL INFORMATION:
; APPLICANT: Orozco Jr., Emil M.
; APPLICANT: Caiami, Perry G.
; APPLICANT: Weng, Zude
; APPLICANT: Tarczynski, Mitchell
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: BB1166 US NA
; CURRENT APPLICATION NUMBER: US/09/697,367
; CURRENT FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/084,529
; PRIOR FILING DATE: 1998-MAY-07
; PRIOR APPLICATION NUMBER: PCT/US99/09865
; PRIOR FILING DATE: 1999-MAY-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1445
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-697-367-19

Query Match 1.3%; Score 44; DB 3; Length 1445;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAAA 3443
|||
DB 1392 TGGCAAA 1435

RESULT 2

US-09-918-909A-19
; Sequence 19, Application US/09918909A
; Patent No. 6756218
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Caiami, Perry G.
; APPLICANT: Orozco Jr., Emil M.
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: BB1166 US CIP
; CURRENT APPLICATION NUMBER: US/09/918,909A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/084,529
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: PCT/US99/09865
; PRIOR FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: 09/697,367
; PRIOR FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1445
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-918-909A-19

Query Match 1.3%; Score 44; DB 4; Length 1445;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAAA 3443
|||
DB 1392 TGGCAAA 1435

RESULT 3

US-09-811-361-19
; Sequence 19, Application US/09811361
; Patent No. 6730821
; GENERAL INFORMATION:
; APPLICANT: Guenther, Catherine
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING
; TITLE OF INVENTION: RETINA-SPECIFIC NUCLEAR RECEPTOR GENE DISRUPTIONS
; FILE REFERENCE: R-125
; CURRENT APPLICATION NUMBER: US/09/811,361
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/190,348
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1530
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-811-361-19

Query Match 1.3%; Score 44; DB 4; Length 1530;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAAA 3443
|||
DB 1478 TGGCAAA 1521

RESULT 4

US-09-500-495A-5
; Sequence 5, Application US/09500495A
; Patent No. 6812382
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Hitz, William D.
; APPLICANT: Lohman, Karin N.
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Plant Nucleic Acids Encoding Chloroplast Division Proteins
; FILE REFERENCE: BB1338 US NA
; CURRENT APPLICATION NUMBER: US/09/500,495A
; CURRENT FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: 60/119,419
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-500-495A-5

Query Match 1.3%; Score 44; DB 4; Length 1560;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3400 TGGCAAA 3443
|||
DB 1491 TGGCAAA 1534

RESULT 5

US-09-345-469-2
; Sequence 2, Application US/09345469
; Patent No. 6369210
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: 22012, A No. 6369210el Human Carboxypeptidase

```

; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 139
; LENGTH: 1720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-227-357-139

Query Match      1.3%; Score 44; DB 3; Length 1720
Best Local Similarity 100.0%; Pred.No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels

QY      3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
          |||
DB      1670 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713

RESULT 7
US-09-614-912-91
; Sequence 91, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Caimi, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/172,959
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/172,946

```

; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 91
; LENGTH: 2311
; TYPE: DNA
; ORGANISM: Glycine max
US-09-614-912-91

Query Match 1.3%; Score 44; DB 4; Length 2311;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||
Db 2255 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2298

RESULT 8

US-09-653-839-9
; Sequence 9, Application US/09653839
; Patent No. 6433153
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
; FILE REFERENCE: LEX-0038-USA
; CURRENT APPLICATION NUMBER: US/09/653,839
; PRIOR FILING DATE: 2000-09-01
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2806
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-653-839-9

Query Match 1.3%; Score 44; DB 3; Length 2806;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||
Db 2709 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2752

RESULT 9

US-10-202-619-9
; Sequence 9, Application US/10202619
; Patent No. 6716614
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6716614el Human Calcium Dependent Proteases,
; FILE REFERENCE: LEX-0391-USA
; CURRENT APPLICATION NUMBER: US/10/202,619
; PRIOR FILING DATE: 2002-07-23
; PRIOR FILING DATE: 1999-09-02
; PRIOR FILING DATE: 1999-09-02
; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2806
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-202-619-9

Query Match 1.3%; Score 44; DB 4; Length 2806;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||
Db 2709 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2752

RESULT 10

US-09-027-137-2
; Sequence 2, Application US/09027137
; Patent No. 6013450
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: CAFI-RELATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,137
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0476 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2852 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT16
; CLONE: 2229466
US-09-027-137-2

Query Match 1.3%; Score 44; DB 3; Length 2852;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
|||
Db 2497 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2540

RESULT 11

US-09-344-441-2
; Sequence 2, Application US/09344441
; Patent No. 6376651
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Corley, Neil C.
; Yue, Henry
; TITLE OF INVENTION: CAP1-RELATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,441
; FILING DATE: 20-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/027,137
; FILING DATE: 1998-02-20
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0476 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2852 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT16
; CLONE: 2229466
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-344-441-2

Query Match 1.3%; Score 44; DB 3; Length 2852;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3400 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 2497 TGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2540

RESULT 12

US-10-160-719A-29
; Sequence 29, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3551)
US-10-160-719A-29

Query Match 1.3%; Score 44; DB 4; Length 3746;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3038 CTCTTCGGCAAGCTCTTCTTCGGCTTCTGGGTATCGTCCACCT 3081
Db 3342 CTCTTCGGCAAGCTCTTCTTCGGCTTCTGGGTATCGTCCACCT 3385

RESULT 13

US-10-160-719A-49
; Sequence 49, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3449)
US-10-160-719A-49

Query Match 1.3%; Score 44; DB 4; Length 3746;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3038 CTCTTCGGCAAGCTCTTCTTCGGCTTCTGGGTATCGTCCACCT 3081
Db 3342 CTCTTCGGCAAGCTCTTCTTCGGCTTCTGGGTATCGTCCACCT 3385

RESULT 14

US-10-160-719A-5
; Sequence 5, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A

RESULT 19
US-08-404-354B-1
; Sequence 1, Application US/08404354B
; Patent No. 5618720
; GENERAL INFORMATION:
; APPLICANT: Ellis, Steven Bradley
; APPLICANT: Williams, Mark E.
; APPLICANT: Harpold, Michael Miller
; APPLICANT: Schwartz, Robert
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404,354B
; FILING DATE: 15-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,083
; FILING DATE: 28-SEP-1994
; APPLICATION NUMBER: US 07/914,231
; FILING DATE: 13-JUL-1992
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 08-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-53192
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 79...5700
; OTHER INFORMATION:
; US-08-404-354B-1
Query Match 1.3%; Score 44; DB 1; Length 5975;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3400 TGGCAA 3443
Db 5928 TGGCAA 5971

RESULT 20
US-08-314-083B-1
; Sequence 1, Application US/08314083B
; Patent No. 5686241
; GENERAL INFORMATION:
; APPLICANT: Ellis, Steven Bradley
; APPLICANT: Williams, Mark E.

; APPLICANT: Harpold, Michael Miller
; APPLICANT: Schwartz, Robert
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/314,083B
; FILING DATE: 28-SEPT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,231
; FILING DATE: 13-JUL-1992
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 08-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-53191
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 79...5700
; OTHER INFORMATION:
; US-08-314-083B-1
Query Match 1.3%; Score 44; DB 1; Length 5975;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3400 TGGCAA 3443
Db 5928 TGGCAA 5971

RESULT 21
US-08-435-675B-1
; Sequence 1, Application US/08435675B
; Patent No. 5710250
; GENERAL INFORMATION:
; APPLICANT: Ellis, Steven Bradley
; APPLICANT: Williams, Mark E.
; APPLICANT: Harpold, Michael Miller
; APPLICANT: Schwartz, Robert
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5975 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 79...5700
OTHER INFORMATION:
US-08-884-599-1

Query Match 1.3%; Score 44; DB 3; Length 5975;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3400 TGGCAA 3443
Db 5928 TGGCAA 5971

RESULT 24

US-09-621-976-16833
Sequence 16833, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 16833
LENGTH: 140
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-16833

Query Match 1.2%; Score 43; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3401 GGCAA 3443
Db 98 GGCAA 140

RESULT 25

US-09-621-976-16834
Sequence 16834, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 16834
LENGTH: 141
TYPE: DNA
ORGANISM: Homo sapiens

US-09-621-976-16834

Query Match 1.2%; Score 43; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3401 GGCAA 3443
Db 98 GGCAA 140

RESULT 26

US-09-621-976-801
Sequence 801, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 801
LENGTH: 194
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 20..193
US-09-621-976-801

Query Match 1.2%; Score 43; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3401 GGCAA 3443
Db 98 GGCAA 140

RESULT 27

US-09-621-976-17670
Sequence 17670, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 17670
LENGTH: 286
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-17670

Query Match 1.2%; Score 43; DB 4; Length 286;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3401 GGCAA 3443
Db 240 GGCAA 282

RESULT 28

```
US-08-721-488-4
; Sequence 4, Application US/08721488
; Patent No. 5965388
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Bowman, Michael
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,488
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 308 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-721-488-4
```

```
Query Match 1.2%; Score 43; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3401 GGCACAAAAA 3443
|||
Db 266 GGCACAAAAA 308
```

```
RESULT 29
US-09-620-405B-465
; Sequence 465, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 465
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-620-405B-465
```

```
Query Match 1.2%; Score 43; DB 4; Length 674;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3401 GGCACAAAAA 3443
|||
Db 547 GGCACAAAAA 589
```

RESULT 30

```
US-09-433-826B-465
; Sequence 465, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 465
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-433-826B-465
```

```
Query Match 1.2%; Score 43; DB 4; Length 674;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3401 GGCACAAAAA 3443
|||
Db 547 GGCACAAAAA 589
```

```
Search completed: March 12, 2005, 07:57:19
Job time : 392 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2005, 05:53:54 ; Search time 1222 Seconds
(without alignments)
16760.046 Million cell updates/sec

Title: US-10-627-132-29

Perfect score: 3443

Sequence: 1 cfcgctgcctcctcgcga.....aaaaaaaaaaaaaaaaaaaa 3443

Scoring table: OLIGO NUC

Gapop 60-0, Gapext 60.0

Searched: 5537552 seqs, 2974263231 residues

Word size : 25

Total number of hits satisfying chosen parameters: 131458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3443	100.0	3443	17	US-10-627-132-29
2	3028	87.9	3028	16	US-10-209-059-29
3	2583	75.0	2687	17	US-10-425-114-833
4	173	5.0	2830	18	US-10-425-115-139789
5	173	5.0	3589	17	US-10-425-114-24930
6	124	3.6	335	18	US-10-425-115-179459
7	116	3.4	999	18	US-10-767-701-10062
8	110	3.2	633	18	US-10-425-115-10141
9	83	2.4	3448	18	US-10-437-963-15093
10	59	1.7	549	18	US-10-767-701-29317
11	58	1.7	488	18	US-10-425-115-139788

1189	9	US-09-900-237-5	Sequence 5, Appli
2351	17	US-10-425-114-2510	Sequence 2510, Ap
2684	18	US-10-425-115-137279	Sequence 137279,
3470	16	US-10-209-059-25	Sequence 25, Appl
3470	17	US-10-627-132-25	Sequence 25, Appl
907	10	US-09-955-999-23	Sequence 23, Appl
1657	18	US-10-437-963-39652	Sequence 39652, A
1795	17	US-10-264-049-155	Sequence 155, App
2218	18	US-10-723-860-7046	Sequence 7046, Ap
431	17	US-10-242-535A-45540	Sequence 45540, A
431	17	US-10-085-783A-45540	Sequence 45540, A
1826	15	US-10-106-698-378	Sequence 378, App
586	18	US-10-425-115-55122	Sequence 55122, A
189	9	US-09-925-298-401	Sequence 401, App
189	14	US-10-102-806-401	Sequence 401, App
358	18	US-10-425-115-102183	Sequence 102183,
711	18	US-10-425-115-11420	Sequence 11420, A
864	18	US-10-425-115-52426	Sequence 52426, A
986	16	US-10-411-224-45	Sequence 105831,
1434	17	US-10-047-021-45	Sequence 45, Appl
1434	14	US-10-227-884-43	Sequence 43, Appl
1434	14	US-10-230-163-43	Sequence 43, Appl
1434	14	US-10-230-338-43	Sequence 43, Appl
1434	14	US-10-218-631-43	Sequence 43, Appl
1434	14	US-10-232-224-43	Sequence 43, Appl
1434	14	US-10-216-159A-43	Sequence 43, Appl
1434	14	US-10-218-849-43	Sequence 43, Appl
1434	14	US-10-227-873-43	Sequence 43, Appl
1434	14	US-10-227-883-43	Sequence 43, Appl
1434	14	US-10-219-076-43	Sequence 43, Appl
1434	14	US-10-230-434-43	Sequence 43, Appl
1434	14	US-10-219-003-43	Sequence 43, Appl
1434	14	US-10-219-075-43	Sequence 43, Appl
1434	14	US-10-219-464-43	Sequence 43, Appl
1434	14	US-10-219-466-43	Sequence 43, Appl
1434	14	US-10-219-471-43	Sequence 43, Appl
1434	14	US-10-219-489-43	Sequence 43, Appl
1434	14	US-10-230-260-43	Sequence 43, Appl
1434	14	US-10-232-231-43	Sequence 43, Appl
1434	14	US-10-232-233-43	Sequence 43, Appl
1434	14	US-10-216-165-43	Sequence 43, Appl
1434	14	US-10-218-956-43	Sequence 43, Appl
1434	14	US-10-219-478-43	Sequence 43, Appl
1434	14	US-10-219-468-43	Sequence 43, Appl
1434	14	US-10-219-536-43	Sequence 43, Appl
1434	14	US-10-233-205-43	Sequence 43, Appl
1434	14	US-10-219-072-43	Sequence 43, Appl
1434	14	US-10-219-470-43	Sequence 43, Appl
1434	14	US-10-219-474-43	Sequence 43, Appl
1434	14	US-10-219-524-43	Sequence 43, Appl
1434	14	US-10-219-528-43	Sequence 43, Appl
1434	14	US-10-232-225-43	Sequence 43, Appl
1434	14	US-10-232-227-43	Sequence 43, Appl
1434	14	US-10-232-229-43	Sequence 43, Appl
1434	14	US-10-232-234-43	Sequence 43, Appl
1434	15	US-10-219-060-43	Sequence 43, Appl
1434	15	US-10-216-160-43	Sequence 43, Appl
1434	15	US-10-216-162-43	Sequence 43, Appl
1434	15	US-10-216-164-43	Sequence 43, Appl
1434	15	US-10-216-167-43	Sequence 43, Appl
1434	15	US-10-219-065-43	Sequence 43, Appl
1434	15	US-10-219-071-43	Sequence 43, Appl
1434	15	US-10-219-074-43	Sequence 43, Appl
1434	15	US-10-219-077-43	Sequence 43, Appl
1434	15	US-10-219-465-43	Sequence 43, Appl

85 45 1.3 1434 15 US-10-219-467-43 Sequence 43, Appl
86 45 1.3 1434 15 US-10-219-469-43 Sequence 43, Appl
87 45 1.3 1434 15 US-10-219-471-43 Sequence 43, Appl
88 45 1.3 1434 15 US-10-219-473-43 Sequence 43, Appl
89 45 1.3 1434 15 US-10-219-476-43 Sequence 43, Appl
90 45 1.3 1434 15 US-10-219-482-43 Sequence 43, Appl
91 45 1.3 1434 15 US-10-227-874-43 Sequence 43, Appl
92 45 1.3 1434 15 US-10-227-876-43 Sequence 43, Appl
93 45 1.3 1434 15 US-10-227-878-43 Sequence 43, Appl
94 45 1.3 1434 15 US-10-229-974-43 Sequence 43, Appl
95 45 1.3 1434 15 US-10-230-024-43 Sequence 43, Appl
96 45 1.3 1434 15 US-10-230-113-43 Sequence 43, Appl
97 45 1.3 1434 15 US-10-230-183-43 Sequence 43, Appl
98 45 1.3 1434 15 US-10-230-234-43 Sequence 43, Appl
99 45 1.3 1434 15 US-10-230-306-43 Sequence 43, Appl
100 45 1.3 1434 15 US-10-230-426-43 Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-10-627-132-29
; Sequence 29, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-29

Query Match 100.0%; Score 3443; DB 17; Length 3443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCGTCCGCTCGCCTCGCAATCGGAATCTGTGAGACACCTGAGGGTGGAGGCCGAG 60
Db 1 CTGCGTCCGCTCGCCTCGCAATCGGAATCTGTGAGACACCTGAGGGTGGAGGCCGAG 60
Qy 61 AGCTAGCTAGCAGCGCGCTCCGCGCGCATGAGAGGCCAGCGCGGCTGTGTGSCCG 120
Db 61 AGCTAGCTAGCAGCGCGCTCCGCGCGCATGAGAGGCCAGCGCGGCTGTGTGSCCG 120
Qy 121 CTCGCAACACCGGAAACGAGTGTGTGATTCGCGGGCCACGAGGACCCCAAGCCGCTGG 180
Db 121 CTCGCAACACCGGAAACGAGTGTGTGATTCGCGGGCCACGAGGACCCCAAGCCGCTGG 180
Qy 181 GCGCTGAGCGGCGAGGTGTGCAGATATCGCGGCGCACGAGTGGGCTCACGTGACCG 240
Db 181 GCGCTGAGCGGCGAGGTGTGCAGATATTCGCGGCGCACGAGTGGGCTCACGTGACCG 240
Qy 241 CGACCTCTTCGTCGCTGCAACAGAGTCGGCTTCCCGGTGTGCGGCTGTGTACAGTA 300
Db 241 CGACCTCTTCGTCGCTGCAACAGAGTCGGCTTCCCGGTGTGCGGCTGTGTACAGTA 300

Qy 301 CGAGCGCCGGAGGGGACCCAGAACTGCCCCCAAGTGTGCAAGCGGCTACAAGCGCCTCAA 360
Db 301 CGAGCGCCGGAGGGGACCCAGAACTGCCCCCAAGTGTGCAAGCGGCTACAAGCGCCTCAA 360
Qy 361 GGGGAGCCCGAGGGTTCGCGGGAGCATGACGAGGAGGACATCGAGACTCGAGCACCAG 420
Db 361 GGGGAGCCCGAGGGTTCGCGGGAGCATGACGAGGAGGACATCGAGACTCGAGCACCAG 420
Qy 421 GTTCAACATCTGACGACGAGAATCAGCAGAGGAGCTGTGAGGGGCAACATCAGAACAGCCA 480
Db 421 GTTCAACATCTGACGACGAGAATCAGCAGAGGAGCTGTGAGGGGCAACATCAGAACAGCCA 480
Qy 481 GATCAACGAGGCGATGCTCCACGCGAGGATGAGTACGAGGAGGGCCCCGACGACGCGGA 540
Db 481 GATCAACGAGGCGATGCTCCACGCGAGGATGAGTACGAGGAGGGCCCCGACGACGCGGA 540
Qy 541 CGGCAACACACCCCGCAGATCCCGCCCATCATACCGGCTCCCGTCCGTCGCGGTGAG 600
Db 541 CGGCAACACACCCCGCAGATCCCGCCCATCATACCGGCTCCCGTCCGTCGCGGTGAG 600
Qy 601 CGGTGAGTTTCCGATTACCAACGGGTATGCGCCACGCGAGGTCTCGTCTTCCCTGCACAA 660
Db 601 CGGTGAGTTTCCGATTACCAACGGGTATGCGCCACGCGAGGTCTCGTCTTCCCTGCACAA 660
Qy 661 GCGCATCCATCCGTACCCCTGTCTGAGCCAGGGATGCCAAGTGGGACGAGAAGAAAGA 720
Db 661 GCGCATCCATCCGTACCCCTGTCTGAGCCAGGGATGCCAAGTGGGACGAGAAGAAAGA 720
Qy 721 AGTGAGCTGGAGGAGAGGATGGAACGACTGGAAGTCCAAAGCAGGGGATCCTCGGCGCG 780
Db 721 AGTGAGCTGGAGGAGAGGATGGAACGACTGGAAGTCCAAAGCAGGGGATCCTCGGCGCG 780
Qy 781 CGCCGATCCCGAAGACATGGAACGCGACTGCGACTGAAACGAGCGAGGCGAGGCGCGT 840
Db 781 CGCCGATCCCGAAGACATGGAACGCGACTGCGACTGAAACGAGCGAGGCGAGGCGCGT 840
Qy 841 GTGAGGAAAGTGTGATCGCTCGAGCAAGGTGAAACCCGTACCGGATGGTGTGATGTT 900
Db 841 GTGAGGAAAGTGTGATCGCTCGAGCAAGGTGAAACCCGTACCGGATGGTGTGATGTT 900
Qy 901 GGGTCTCGTGTGTGCTCGCCTTCTTCTCGGTACCGTATCTCTGCAACCCCGTCCCGACGC 960
Db 901 GGGTCTCGTGTGTGCTCGCCTTCTTCTCGGTACCGTATCTCTGCAACCCCGTCCCGACGC 960
Qy 961 CATCGGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGTTTCGCCATCTCTCTGGATCT 1020
Db 961 CATCGGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGTTTCGCCATCTCTCTGGATCT 1020
Qy 1021 CGACAGTTTCCCAAGTGGTTCCTCATCGACCGGAGAGTACCTCGACCGGCTCTCCCT 1080
Db 1021 CGACAGTTTCCCAAGTGGTTCCTCATCGACCGGAGAGTACCTCGACCGGCTCTCCCT 1080
Qy 1081 CAGGTACGAGAGGAGGAGCGCTGCTGTGTCGGGCTGGACCTGTTCTGTGAGCAC 1140
Db 1081 CAGGTACGAGAGGAGGAGCGCTGCTGTGTCGGGCTGGACCTGTTCTGTGAGCAC 1140
Qy 1141 GGTGACCCCGCTCAAGGAGCCCGCTGCTGAGCCGCAACACCGTGTCTCTCATCTCTCGC 1200
Db 1141 GGTGACCCCGCTCAAGGAGCCCGCTGCTGAGCCGCAACACCGTGTCTCTCATCTCTCGC 1200
Qy 1201 CTTAGACTACCCGCTGGACAAAGTCTCTGTCTGTCAGTCTCCGACGAGCGGCTGTGATGT 1260
Db 1201 CTTAGACTACCCGCTGGACAAAGTCTCTGTCTGTCAGTCTCCGACGAGCGGCTGTGATGT 1260
Qy 1261 GACGTTTCGAGTCTGTCGAGACCGGCTGAGTTCGCGGCAAGTGGGTGCCCTTCTGCAA 1320
Db 1261 GACGTTTCGAGTCTGTCGAGACCGGCTGAGTTCGCGGCAAGTGGGTGCCCTTCTGCAA 1320
Qy 1321 GAAGTTTCGCACTGAGACCCCGGCTGAGTTCCTCTCTGCTCAAGGTGAGTACCT 1380
Db 1321 GAAGTTTCGCACTGAGACCCCGGCTGAGTTCCTCTCTGCTCAAGGTGAGTACCT 1380
Qy 1381 CAAAGCAAGGTGACGCCCACTTCTGTGAGGAGCGCGGCCATGAAGAGAGATGTA 1440

Db 1381 CAAAGCAAGGTGACGCCACCTTCGTGACGAGCGCGGCCCATGAGAGAGATGA 1440
Qy |||||
Db 1441 GGAGTTCAAGTTCGGATCAACCGGCTGTGGTCCCAAGGCCATGAAGGTGCCGGCAGAGGG 1500
Qy |||||
Db 1441 GGAGTTCAAGTTCGGATCAACCGGCTGTGGTCCCAAGGCCATGAAGGTGCCGGCAGAGGG 1500
Qy |||||
Db 1501 GTGATCATGAGAGCGGACCGCGTGGTCCCGGGAAACAAACCCCGGCGAT 1560
Qy |||||
Db 1501 GTGATCATGAGAGCGGACCGCGTGGTCCCGGGAAACAAACCCCGGCGAT 1560
Qy |||||
Db 1561 GATCCAGGTGTTCTTGGGCCACAGCGCGGCGACGACACAGAGGGCAACAGCTGCCCG 1620
Qy |||||
Db 1561 GATCCAGGTGTTCTTGGGCCACAGCGCGGCGACGACACAGAGGGCAACAGCTGCCCG 1620
Qy |||||
Db 1621 CCTCGTGTACGTCTCCGTGAGAGCGCCGGGATTCAGCACCAAGAGCGCGCGC 1680
Qy |||||
Db 1621 CCTCGTGTACGTCTCCGTGAGAGCGCCGGGATTCAGCACCAAGAGCGCGCGC 1680
Qy |||||
Db 1681 CATGAACGCTCTGATTCGGCTCTCGCGTCTGACCAACCGCGCATTCATGCTCAACTT 1740
Qy |||||
Db 1681 CATGAACGCTCTGATTCGGCTCTCGCGTCTGACCAACCGCGCATTCATGCTCAACTT 1740
Qy |||||
Db 1741 GGACTGTGATCACTACATCAACAACAGAGGCCATCCGGAGGCCATGCTTCTCAT 1800
Qy |||||
Db 1741 GGACTGTGATCACTACATCAACAACAGAGGCCATCCGGAGGCCATGCTTCTCAT 1800
Qy |||||
Db 1801 GGACCTCAGGTCCGGCGGAGGTCTGCTAGTTCAGTTCGCGAGGTTCCAGCGCAT 1860
Qy |||||
Db 1801 GGACCTCAGGTCCGGCGGAGGTCTGCTAGTTCAGTTCGCGAGGTTCCAGCGCAT 1860
Qy |||||
Db 1861 CGAGTGCACGACCGATACGCTAACAGAGAACACCGTCTTCTTGACATCAACATGAAGGG 1920
Qy |||||
Db 1861 CGAGTGCACGACCGATACGCTAACAGAGAACACCGTCTTCTTGACATCAACATGAAGGG 1920
Qy |||||
Db 1921 GCTGACGGCATCCAAAGGCCGGGTACGTCCGGACAGGGTGGTTCGGGCGCCAGGC 1980
Qy |||||
Db 1921 GCTGACGGCATCCAAAGGCCGGGTACGTCCGGACAGGGTGGTTCGGGCGCCAGGC 1980
Qy |||||
Db 1981 GCTCTACGGCTACAAACCTCTCCAAAGGACCCAAAGGCCCAAGATGGTGACCTCCGACTG 2040
Qy |||||
Db 1981 GCTCTACGGCTACAAACCTCTCCAAAGGACCCAAAGGCCCAAGATGGTGACCTCCGACTG 2040
Qy |||||
Db 2041 CTGCGCGTGTTCGGCGCGAAGAGCGGAAACAGCGCAAGGACGGGTGCGCGAGGCAC 2100
Qy |||||
Db 2041 CTGCGCGTGTTCGGCGCGAAGAGCGGAAACAGCGCAAGGACGGGTGCGCGAGGCAC 2100
Qy |||||
Db 2101 CGCTGATATGGAGTAGATAGCGACAGGAGATGCTCATGTCCACATGAATTCGAGAA 2160
Qy |||||
Db 2101 CGCTGATATGGAGTAGATAGCGACAGGAGATGCTCATGTCCACATGAATTCGAGAA 2160
Qy |||||
Db 2161 GCGGTTCCGGCAGTCCGCGGCTTCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2220
Qy |||||
Db 2161 GCGGTTCCGGCAGTCCGCGGCTTCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2220
Qy |||||
Db 2221 TCCTTCGTGAGCCCGCGCGTCTCAAGAGGCCATCATGTCAATCAAGTCCGCGCTA 2280
Qy |||||
Db 2221 TCCTTCGTGAGCCCGCGCGTCTCAAGAGGCCATCATGTCAATCAAGTCCGCGCTA 2280
Qy |||||
Db 2281 CGAGGACAGACCGACTGGGGCTGGAGCTGGGTGGATCTACGGGTGATCAGCGAGGA 2340
Qy |||||
Db 2281 CGAGGACAGACCGACTGGGGCTGGAGCTGGGTGGATCTACGGGTGATCAGCGAGGA 2340
Qy |||||
Db 2341 CATCTGACGGGGTTCAAGATGCACTGCGCGGGTGGCGTCCGTGTACTGCAATCCGGA 2400
Qy |||||
Db 2341 CATCTGACGGGGTTCAAGATGCACTGCGCGGGTGGCGTCCGTGTACTGCAATCCGGA 2400
Qy |||||
Db 2401 GCGGGCGCGGTTCAAGGGTCCGCGCGATCAATCTATCGAGCGTCTCAACAGGTGCT 2460
Qy |||||
Db 2401 GCGGGCGCGGTTCAAGGGTCCGCGCGATCAATCTATCGAGCGTCTCAACAGGTGCT 2460
Qy |||||
Db 2461 CCGGTGGCGCTGGGGTCCGTGAGATCTTCTCAGCGCGGACAGCGCCCTGCTGTACGG 2520
Qy |||||

Db 2461 CCGGTGGCGCTGGGTCCGTGAGATCTTCTTCAAGCGGCGACAGCCCCCTGCTGTACGG 2520
Qy |||||
Db 2521 CTACAAGAACGGCAACCTCAAGTGGTGGAGCGCTTCGCTCAATCAACACCAACCATCTA 2580
Qy |||||
Db 2521 CTACAAGAACGGCAACCTCAAGTGGTGGAGCGCTTCGCTCAATCAACACCAACCATCTA 2580
Qy |||||
Db 2581 CCGCTTCACTCCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Qy |||||
Db 2581 CCGCTTCACTCCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Qy |||||
Db 2641 CCGCAAGTTCAATCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Qy |||||
Db 2641 CCGCAAGTTCAATCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Qy |||||
Db 2701 CATGTCCATCTTCGCGACGGGATCTCGGATGCGGTGAGCGGGGTGAGCATCGAGGA 2760
Qy |||||
Db 2701 CATGTCCATCTTCGCGACGGGATCTCGGATGCGGTGAGCGGGGTGAGCATCGAGGA 2760
Qy |||||
Db 2761 GTGGTGGAGAAAGAGCAGTTCGGGTGCTGCGGGGGGTGCTGCGCGCATCTCTTCGCGGT 2820
Qy |||||
Db 2761 GTGGTGGAGAAAGAGCAGTTCGGGTGCTGCGGGGGGTGCTGCGCGCATCTCTTCGCGGT 2820
Qy |||||
Db 2821 CGTGCAGGGCTCTCAAGGTCTCGCGGGATCGACACCACTTCAACCGTCACTTCAA 2880
Qy |||||
Db 2821 CGTGCAGGGCTCTCAAGGTCTCGCGGGATCGACACCACTTCAACCGTCACTTCAA 2880
Qy |||||
Db 2881 GGGCACGGCGACGAGGACGACGAGTTCGCCGAGCTCTACGCCCTTCAAGTGGACACGCT 2940
Qy |||||
Db 2881 GGGCACGGCGACGAGGACGACGAGTTCGCCGAGCTCTACGCCCTTCAAGTGGACACGCT 2940
Qy |||||
Db 2941 CCTCATCCCGCCACACGCTGCTCATTAAGTCAATGCTGCTGCGGTGCTGCGCGCATCTC 3000
Qy |||||
Db 2941 CCTCATCCCGCCACACGCTGCTCATTAAGTCAATGCTGCTGCGGTGCTGCGCGCATCTC 3000
Qy |||||
Db 3001 CGACGCCATCAACAAAGGGTACGAGTCTCGGGGGCCCTCTCGGCAAGCTCTTCTTCGCG 3060
Qy |||||
Db 3001 CGACGCCATCAACAAAGGGTACGAGTCTCGGGGGCCCTCTCGGCAAGCTCTTCTTCGCG 3060
Qy |||||
Db 3061 CTTCTGGGTCACTCGTCCACCTCTACCGTTCCTCAAGGGGCTCATGGGGCGCCAGAACAG 3120
Qy |||||
Db 3061 CTTCTGGGTCACTCGTCCACCTCTACCGTTCCTCAAGGGGCTCATGGGGCGCCAGAACAG 3120
Qy |||||
Db 3121 GAGCGCCACGTTGTTGTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180
Qy |||||
Db 3121 GAGCGCCACGTTGTTGTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180
Qy |||||
Db 3181 GGTGAGATCGACCTCTTCAATCGTCAGGACCAAGGGCCCGGACGTCAGGACAGTGTGGCAT 3240
Qy |||||
Db 3181 GGTGAGATCGACCTCTTCAATCGTCAGGACCAAGGGCCCGGACGTCAGGACAGTGTGGCAT 3240
Qy |||||
Db 3241 CAATTCGAGCTGTTTATTAAGGTTCAAAATTCGAGCTTGTGCAATAGGAGAGAAAAA 3300
Qy |||||
Db 3241 CAATTCGAGCTGTTTATTAAGGTTCAAAATTCGAGCTTGTGCAATAGGAGAGAAAAA 3300
Qy |||||
Db 3301 ACAATTTAGAAATTTGTAAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3360
Qy |||||
Db 3301 ACAATTTAGAAATTTGTAAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3360
Qy |||||
Db 3361 GAGGAGAAATTAAGAAAGGACAAAGTTGATTGTTAAATGGCAAAAAA 3420
Qy |||||
Db 3361 GAGGAGAAATTAAGAAAGGACAAAGTTGATTGTTAAATGGCAAAAAA 3420
Qy |||||
Db 3421 AAAAAAAAAAAAAAAAAAAAAA 3443
Qy |||||
Db 3421 AAAAAAAAAAAAAAAAAAAAAA 3443

RESULT 2

US-10-059-29

; Sequence 29, Application US/10209059

; Publication No. US200301638A1

; GENERAL INFORMATION:

; APPLICANT: Dhugga, Kanwarpal S.

```
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3028
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-29

Query Match      87.9%; Score 3028; DB 16; Length 3028;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3028; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 CACGAGTTCAACATCGACGACGAGATCAGCAGAGCAGCTGAGGAGGCAACATGCAGAAC 475
DB 1 CACGAGTTCAACATCGACGACGAGATCAGCAGAGGAGCTGGAGGAGCAACATGCAGAAC 60

QY 476 AGCCAGATCACCGAGGAGTGTGCGACGGCAGAGTACGAGGGGGCCCGCAGCAC 535
DB 61 AGCCAGATCACCGAGGAGTGTGCGACGGCAGAGTACGAGGGAGGGGCCCGCAGCAC 120

QY 536 GCGGACGGCAACACACCCGCGAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGCG 595
DB 121 GCGGACGGCAACACACCCGCGAGATCCCGCCCATCATCACCGGCTCCCGCTCCGTCGCG 180

QY 596 GTGAGCGGTGAGTTTCGATTAACAACGGGTATGGCCAGCGGAGGTCTGCTTTCCTTG 655
DB 181 GTGAGCGGTGAGTTTCGATTAACAACGGGTATGGCCAGCGGAGGTCTGCTTTCCTTG 240

QY 656 CACAAGCGCATCCATCGTACCTGTCTGTGAGCCAGGAGTCCCAAGTGGGACGAGAG 715
DB 241 CACAAGCGCATCCATCGTACCTGTCTGTGAGCCAGGAGTCCCAAGTGGGACGAGAG 300

QY 716 AAAGAAGTGAAGTGGAGGAGAGATGGAACGACTGGAAGTCCAAAGCAGGAGCATCTCTCGG 775
DB 301 AAAGAAGTGAAGTGGAGGAGAGATGGAACGACTGGAAGTCCAAAGCAGGAGCATCTCTCGG 360

QY 776 GCGGCGCCGATCCCGAAGACATGAGACCGCGACGCTGGCACTGAACGACGAGGCGAGGCGAG 835
DB 361 GCGGCGCCGATCCCGAAGACATGAGACCGCGACGCTGGCACTGAACGACGAGGCGAGGCGAG 420

QY 836 CCGCTGTCGAGGAGGTGTCGATCGCTGAGCAAGAGTGAACCGGTACCGGATGGGTGATC 895
DB 421 CCGCTGTCGAGGAGGTGTCGATCGCTGAGCAAGAGTGAACCGGTACCGGATGGGTGATC 480

QY 896 GTGCTGCGTCTCGTGTGCTCGCTTCTTCTCGGTACCGGTATCTCTGCAACCCCGCTCCCG 955
DB 481 GTGCTGCGTCTCGTGTGCTCGCTTCTTCTCGGTACCGGTATCTCTGCAACCCCGCTCCCG 540

QY 956 GAGCCCATCGGGCTGTGGCTGCTTCATCATCTGCGAGATCTGTTGGCCATCTCTCTGG 1015
DB 541 GAGCCCATCGGGCTGTGGCTGCTTCATCATCTGCGAGATCTGTTGGCCATCTCTCTGG 600

QY 1016 ATCTCGACAGTTCCTCCAGTGTTCCTCCATCGACCGGAGACGTACCTCGACCGCTC 1075
DB 601 ATCTCGACAGTTCCTCCAGTGTTCCTCCATCGACCGGAGACGTACCTCGACCGCTC 660

QY 1076 TCCTCAGGTACGAGAGGAGGAGCGCTGCTGCTGTGCGGCGGTGAGACCTGTTCTG 1135
DB 661 TCCTCAGGTACGAGAGGAGGAGCGCTGCTGCTGTGCGGCGGTGAGACCTGTTCTG 720

QY 1136 AGCAGGTGACCCGCTCAAGGAGCGCCGCTGTGTGACCCGCCAACACCGTGTCTCTCATC 1195
```

```
DB 721 AGCAGGTGAGCCCGCTCAAGGAGCGCGCTGTGTGACCCCAACACCGTGTCTCTCATC 780
QY 1196 CTCGCCGTAGACTACCCCGTGGACAAGGTCTCTCTGTCTGTCTCTCCGACGACGCGGTGCG 1255
DB 781 CTCGCCGTAGACTACCCCGTGGACAAGGTCTCTCTGTCTGTCTCTCCGACGACGCGGTGCG 840
QY 1256 ATGTGACGTTTCGAGTCTGTCGAGACGCGCGAGTTCGCGCGCAAGTGGGTGCCCTTC 1315
DB 841 ATGTGACGTTTCGAGTCTGTCGAGACGCGCGAGTTCGCGCGCAAGTGGGTGCCCTTC 900
QY 1316 TGCAGAAGTTTCGGCATCGAGCCCGCGCGCCCGGAGTTCTACTTCTCGCTCAAGTTCGAC 1375
DB 901 TGCAGAAGTTTCGGCATCGAGCCCGCGCGCCCGGAGTTCTACTTCTCGCTCAAGTTCGAC 960
QY 1376 TACCTCAAGGACAAGGTGAGCCCACTTCTGTGAGGAGCGCGCGCATGAAGAGAGAG 1435
DB 961 TACCTCAAGGACAAGGTGAGCCCACTTCTGTGAGGAGCGCGCGCATGAAGAGAGAG 1020
QY 1436 TATGAGGAGTTTCAAGTCCGGATCAACGCGCTGTGTGGCCAAAGGCCATGAAGTTCGCGCA 1495
DB 1021 TATGAGGAGTTTCAAGTCCGGATCAACGCGCTGTGTGGCCAAAGGCCATGAAGTTCGCGCA 1080
QY 1496 GAGGGGTGATCATGAAGACGCGACCGCGTGGCCCGGGAAACAACCCGCGACCAACCCC 1555
DB 1081 GAGGGGTGATCATGAAGACGCGACCGCGTGGCCCGGGAAACAACCCGCGACCAACCCC 1140
QY 1556 GGCATGATCACAGGTTCCTGGGCCACAGCGCGGCGCACACACCGAGGCAACGAGCTG 1615
DB 1141 GGCATGATCACAGGTTCCTGGGCCACAGCGCGGCGCACACACCGAGGCAACGAGCTG 1200
QY 1616 CCGCGCTCTGTGTACGTCTCCGTGAGAACGCGCGGATTCAGCAACAACAAGAGGCC 1675
DB 1201 CCGCGCTCTGTGTACGTCTCCGTGAGAACGCGCGGATTCAGCAACAACAAGAGGCC 1260
QY 1676 GGGCCCATGAACGCTCTGATTCGGCTTCGGCGTCTGACCAACGCGGCATTCATGCTC 1735
DB 1261 GGGCCCATGAACGCTCTGATTCGGCTTCGGCGTCTGACCAACGCGGCATTCATGCTC 1320
QY 1736 AACTTGGAGTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCAATGTGCTTC 1795
DB 1321 AACTTGGAGTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGCAATGTGCTTC 1380
QY 1796 CTCATGAGCCTCAGGTCCGCGCGAGAGTCTGCTACGTTTCAGTTCCTCCGAGAGTTCGAC 1855
DB 1381 CTCATGAGCCTCAGGTCCGCGCGAGAGTCTGCTACGTTTCAGTTCCTCCGAGAGTTCGAC 1440
QY 1856 GGCATCGAGTGCACGACCGATACGCTAACAGGAAACACCGTCTTCTTCGACATCAACATG 1915
DB 1441 GGCATCGAGTGCACGACCGATACGCTAACAGGAAACACCGTCTTCTTCGACATCAACATG 1500
QY 1916 AAGGGGCTGGACGGCATCCAAAGCCCGGTGTACGTCCGAGACAGGGTTCGTTCCGGCGC 1975
DB 1501 AAGGGGCTGGACGGCATCCAAAGCCCGGTGTACGTCCGAGACAGGGTTCGTTTCGGCGC 1560
QY 1976 CAGCGGCTCTACGGCTACAACCTCCAAAGGACCCAAAGAGGCCCAAGATGTTGACCTGC 2035
DB 1561 CAGCGGCTCTACGGCTACAACCTCCCAAGGACCCAAAGAGGCCCAAGATGTTGACCTGC 1620
QY 2036 GACTGTGCGCTGTCTTCGGCGCAAGAGCGGAAACACGCCAAGACCGGGTTCGCGGAG 2095
DB 1621 GACTGTGCGCTGTCTTCGGCGCAAGAGCGGAAACACGCCAAGACCGGGTTCGCGGAG 1680
QY 2096 GGCACCGCTGTATATGGAGTAGATAGCGACAAGAGATGCTCATGTGCCACATGAACCTTC 2155
DB 1681 GGCACCGCTGTATATGGAGTAGATAGCGACAAGAGATGCTCATGTGCCACATGAACCTTC 1740
QY 2156 GAGAAGCGGTTCGGCGAGTCCGCGGCTTCGTCACGTCGACGCTGATGAGGAGGAGCGGC 2215
DB 1741 GAGAAGCGGTTCGGCGAGTCCGCGGCTTCGTCACGTCGACGCTGATGAGGAGGAGCGGC 1800
QY 2216 GTCCCTCTCTGTGAGAGCCCGCGCTCCTCAAGGAGGCCATTCATGTCTCATGCTGCTC 2275
```

Db 1801 GTCCCTCCTTCTCGAGCGCCCGCGCTCCTCAAGGAGGCCATCCATGTCATCATGCTGC 1860
Qy 2276 GGCTACGAGGACAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTGATCAACG 2335
Db 1861 GGCTACGAGGACAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTGATCAACG 1920
Qy 2336 GAGGACATCTGAGCGGGTTCAGATGCACTGCGCGGGTGGCGCTCGGTGACTGATG 2395
Db 1921 GAGGACATCTGAGCGGGTTCAGATGCACTGCGCGGGTGGCGCTCGGTGACTGATG 1980
Qy 2396 CCGAAGCGGGCGCGGTTCAAGGGGTGCGCGCCGATCAATCTATCGGACCGCTCTCAACGAG 2455
Db 1981 CCGAAGCGGGCGGTTCAAGGGGTGCGCGCCGATCAATCTATCGGACCGCTCTCAACGAG 2040
Qy 2456 GTGCTCCGGTGGCGCTGGGGTCCGTGAGATCTTTCAGCGCGGACAGCCCCCTGCTG 2515
Db 2041 GTGCTCCGGTGGCGCTGGGGTCCGTGAGATCTTTCAGCGCGGACAGCCCCCTGCTG 2100
Qy 2516 TAGGGCTACAGAACGCGCACTCAAGTGGCTGGAGCGCTTCGGCTACATCAACACCAACC 2575
Db 2101 TAGGGCTACAGAACGCGCACTCAAGTGGCTGGAGCGCTTCGGCTACATCAACACCAACC 2160
Qy 2576 ATCTACCCCTTCACTCGCTCCGCTGCTGCGCTACTGACCCCTCCCGCGCTCTGCTC 2635
Db 2161 ATCTACCCCTTCACTCGCTCCGCTGCTGCGCTACTGACCCCTCCCGCGCTCTGCTC 2220
Qy 2636 CTCACCGGCAAGTTTCATCATGCGCGCTGATAGCAGTTTCGCGACCTCTTCTTCATCGCC 2695
Db 2221 CTCACCGGCAAGTTTCATCATGCGCGCTGATAGCAGTTTCGCGACCTCTTCTTCATCGCC 2280
Qy 2696 CTCTTCATGTCATCTTCGCGACCGGCACTCTGAGATGCGGTGGAGCGGGGTGAGCATC 2755
Db 2281 CTCTTCATGTCATCTTCGCGACCGGCACTCTGAGATGCGGTGGAGCGGGGTGAGCATC 2340
Qy 2756 GAGGAGTGGTGGAGGACGAGCAGTTCTGGTGCATCGCGCGGTGTCGCGCATCTCTTC 2815
Db 2341 GAGGAGTGGTGGAGGACGAGCAGTTCTGGTGCATCGCGCGGTGTCGCGCATCTCTTC 2400
Qy 2816 GCCGTGTCAGGGCTGCTCAAGTCTCTCGCGGGATCGACACCAACTTCACCGTCACCC 2875
Db 2401 GCCGTGTCAGGGCTGCTCAAGTCTCTCGCGGGATCGACACCAACTTCACCGTCACCC 2460
Qy 2876 TCCAGGCGCACCGCGGACGAGGACGAGTTTCGCGAGCTCTACGGCTTCAGTGGAC 2935
Db 2461 TCCAGGCGCACCGCGGACGAGGACGAGTTTCGCGAGCTCTACGGCTTCAGTGGAC 2520
Qy 2936 ACGCTCTCTATCCCGCCACACACGCTGCTCATCAATTAACGTCATCGCGGTGCTGGCGGC 2995
Db 2521 ACGCTCTCTATCCCGCCACACACGCTGCTCATCAATTAACGTCATCGCGGTGCTGGCGGC 2580
Qy 2996 ATCTCCGACGCGCATCAACACGGGTACAGTCTCTGGGGGCCCTCTTCGCGCAAGCTCTTC 3055
Db 2581 ATCTCCGACGCGCATCAACACGGGTACAGTCTCTGGGGGCCCTCTTCGCGCAAGCTCTTC 2640
Qy 3056 TTGCGCTTCTGGGTGATGTCACCTCTACCGGTTCCTCAAGGGGCTCATGGGGGCGCAG 3115
Db 2641 TTGCGCTTCTGGGTGATGTCACCTCTACCGGTTCCTCAAGGGGCTCATGGGGGCGCAG 2700
Qy 3116 AACAGGACGCCACCGGTGTTGTGTCATCTGGTCCATCTGCTGGCGCTCCATCTTCTCCCTG 3175
Db 2701 AACAGGACGCCACCGGTGTTGTGTCATCTGGTCCATCTGCTGGCGCTCCATCTTCTCCCTG 2760
Qy 3176 CTCTGGGTGAGGATCGAGCCCTTTTCATCGTCAAGAACCAAGGCGCCGAGCGTCAGGAGTGT 3235
Db 2761 CTCTGGGTGAGGATCGAGCCCTTTTCATCGTCAAGAACCAAGGCGCCGAGCGTCAGGAGTGT 2820
Qy 3236 GGCATCAATGCTGAGCTGTTTATTAAGGTTCAAAATTCCTGGAGCTTGTGTCATAGGAGA 3295
Db 2821 GGCATCAATGCTGAGCTGTTTATTAAGGTTCAAAATTCCTGGAGCTTGTGTCATAGGAGA 2880
Qy 3296 AAAAAACAAATTTAGAAATTTTGAAGTTGTTGTCTGTAATGTTATGTTATGCCAGAAAT 3355
Db 2881 AAAAAACAAATTTTGAAGTTGTTGTCTGTAATGTTATGTTATGCCAGAAAT 2940

Qy 3356 TGTGCGACGAGGAATTGAACAAAGGACAAGGTTTGTATTTGTTAAATGCAAAAAA 3415
Db 2941 TGTGCGACGAGGAATTGAACAAAGGACAAGGTTTGTATTTGTTAAATGCAAAAAA 3000
Qy 3416 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 3001 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3028

RESULT 3

US-10-425-114-833
; Sequence 833, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 833
; LENGTH: 2687
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700093978_FLI
US-10-425-114-833

Query Match 75.0%; Score 2583; DB 17; Length 2687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 773 GCGCGGCGGCCCATCCGAAAGACATGCGCGAGTGGCACTGAAACGACGAGCGCAGG 832
Db 1 GCGCGGCGGCCCATCCGAGACATGCGCGAGTGGCACTGAAACGACGAGCGCAGG 60
Qy 833 CAGCGCTGTCGAGGAAGGTGTCGATCGCGTGCAGCAAGGTGAACCGGTAACCGATGGT 892
Db 61 CAGCGCTGTCGAGGAAGGTGTCGATCGCGTGCAGCAAGGTGAACCGGTAACCGATGGT 120
Qy 893 ATCGTGGTGGCTCTCGTTGGCTCGCTTCTTCTCGGTACCGTATCCTGCAACCGCTC 952
Db 121 ATCGTGGTGGCTCTCGTTGGCTCGCTTCTTCTCGGTACCGTATCCTGCAACCGCTC 180
Qy 953 CCGGACGCCATCGGCTGCGCTCGTCTCCATCATCTCGAGATCTGGTTGCGCATCTCC 1012
Db 181 CCGGACGCCATCGGCTGCGCTCGTCTCCATCATCTCGAGATCTGGTTGCGCATCTCC 240
Qy 1013 TGGATCCTCGACAGTTCGCCAAGTGGTTCGCCATCGACCGAGAGCTACCTCGACCGC 1072
Db 241 TGGATCCTCGACAGTTCGCCAAGTGGTTCGCCATCGACCGAGAGCTACCTCGACCGC 300
Qy 1073 CTCTCCTCAGGTACGAGGAGGAGGAGCGCTCGTCTGCTGCGCGGTGGACCTGTTTC 1132
Db 301 CTCTCCTCAGGTACGAGGAGGAGGAGGAGCGCTCGTCTGCTGCGCGGTGGACCTGTTTC 360
Qy 1133 GTGAGCAGGTGGACCGCTCAAGGAGCGCGCTGGTGCACCGCAACCGCTGCTCTCC 1192
Db 361 GTGAGCAGGTGGACCGCTCAAGGAGCGCGCTGGTGCACCGCAACCGCTGCTCTCC 420
Qy 1193 ATCTCGCGGTAGACTACCGCGTGGACAGGTCTCTCTGCTACGTTCTCGACGAGCGCGC 1252
Db 421 ATCTCGCGGTAGACTACCGCGTGGACAGGTCTCTCTGCTACGTTCTCGACGAGCGCGC 480
Qy 1253 TCGATGCTGACGTTGAGTCTGTCGAGACCGCGCGAGTTCGCGCGCAAGTGGGTGCCCC 1312

Db 481 TCGATGCTGAGTTGAGTGTCTGTGCGAGACGGCCGAGTTTCGCGCGCAAGTGGGTGCC 540
Qy 1313 TTCTGCAAGAAGTTCCGCAATCGAGCCCGCGCCCGAGTTCTTAATCTTCGCTCAAGGTC 1372
Db 541 TTCTGCAAGAAGTTCCGCAATCGAGCCCGCGCCCGAGTTCTTAATCTTCGCTCAAGGTC 600
Qy 1373 GACTACTCAAGACAAAGGTGAGCCCACTTCGTGACAGAGCGCGCCCATGAAGAGA 1432
Db 601 GACTACTCAAGACAAAGGTGAGCCCACTTCGTGACAGAGCGCGCCCATGAAGAGA 660
Qy 1433 GAGTATGAGAGTTCAAGGTTCCGATCAACCGCTGTGTGSCCAAGGCCATGAAGGTCCG 1492
Db 661 GAGTATGAGAGTTCAAGGTTCCGATCAACCGCTGTGTGSCCAAGGCCATGAAGGTCCG 720
Qy 1493 GCAGAGGGTGGATCATGAAGAGCGGCAACCGCTGTGCGGGAAACAACCCGCGACAC 1552
Db 721 GCAGAGGGTGGATCATGAAGAGCGGCAACCGCTGTGCGGGAAACAACCCGCGACAC 780
Qy 1553 CCCGGCATGATCAGGTGTTCTTGGGCAAGCGGCGGCCACGACACCGAGGCAACGAG 1612
Db 781 CCCGGCATGATCAGGTGTTCTTGGGCAAGCGGCGGCCACGACACCGAGGCAACGAG 840
Qy 1613 CTGCCCCGCTCGTGTACGTTCTCCGTTGAGAGCGCCCGGGATTCAGACCAACAAGAG 1672
Db 841 CTGCCCCGCTCGTGTACGTTCTCCGTTGAGAGCGCCCGGGATTCAGACCAACAAGAG 900
Qy 1673 GCGGGGCCATGAACGCTCTGATTCTCGCTCTCGCGCTGTGACCAACCGCCCATTCATG 1732
Db 901 GCGGGGCCATGAACGCTCTGATTCTCGCTCTCGCGCTGTGACCAACCGCCCATTCATG 960
Qy 1733 CTCAACTTGACCTGTGATCATCATCAACAAGCAAGGCCATTCGCGAGGCCATGTGC 1792
Db 961 CTCAACTTGACCTGTGATCATCATCAACAAGCAAGGCCATTCGCGAGGCCATGTGC 1020
Qy 1793 TTCTCATGACCCCTCAGTCCGCGGAGGTCTGTAAGTTTCCGAGAGGTTTC 1852
Db 1021 TTCTCATGACCCCTCAGTCCGCGGAGGTCTGTAAGTTTCCGAGAGGTTTC 1080
Qy 1853 GACGGCATGACGTGACGACCGATACGCTAAACAGGAACACCGCTCTTTCGACATCAAC 1912
Db 1081 GACGGCATGACGTGACGACCGATACGCTAAACAGGAACACCGCTCTTTCGACATCAAC 1140
Qy 1913 ATGAAGGGCTGACGGCATCAAGGCCGGTGTACGTGCGGACAGGGTGTGTTCCGG 1972
Db 1141 ATGAAGGGCTGACGGCATCAAGGCCGGTGTACGTGCGGACAGGGTGTGTTCCGG 1200
Qy 1973 CGCAGCGCTCTACGGCTAACACCTCCCAAGGACCCCAAGAGGCCAAGATGTTGAC 2032
Db 1201 CGCAGCGCTCTACGGCTAACACCTCCCAAGGACCCCAAGAGGCCAAGATGTTGAC 1260
Qy 2033 TCGGACTGTGCGCGTCTTCGCGCGCAAGAGCGGAAACACGCCAAGGACGGGCTGCCG 2092
Db 1261 TCGGACTGTGCGCGTCTTCGCGCGCAAGAGCGGAAACACGCCAAGGACGGGCTGCCG 1320
Qy 2093 GAGGGACCGCTGATATGGAGTAGATAGCGAACAGAGATGTCTATGTCCCAATGAAC 2152
Db 1321 GAGGGACCGCTGATATGGAGTAGATAGCGAACAGAGATGTCTATGTCCCAATGAAC 1380
Qy 2153 TTGAGAGCGGTTCCGGCAGTCCGGCGTTCGTCAAGTCCGAGCTGTATGGAGGAGGC 2212
Db 1381 TTGAGAGCGGTTCCGGCAGTCCGGCGTTCGTCAAGTCCGAGCTGTATGGAGGAGGC 1440
Qy 2213 GCGGTCTCTTCTGTCGAGCCCGCGCTCTCAAGAGGCCATCCATGTATCATCAGC 2272
Db 1441 GCGGTCTCTTCTGTCGAGCCCGCGCTCTCAAGAGGCCATCCATGTATCATCAGC 1500
Qy 2273 TGGCGCTACAGGACAAAGCCGACTGGGGCTGGAGCTGGGGTGGATCTTACGGGTGCATC 2332
Db 1501 TGGCGCTACAGGACAAAGCCGACTGGGGCTGGAGCTGGGGTGGATCTTACGGGTGCATC 1560
Qy 2333 ACGGAGGACATCTGACGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTGTACTGC 2392
Db 1561 ACGGAGGACATCTGACGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTGTACTGC 1620

RESULT 4

US-10-425-115-139789

; Sequence 139789, Application US/10425115

Qy 2393 ATGCCGAAGCGGCGCGGTTCAAGGGGTGCGCGCGCATCAATCTATCGAACGCTCTCAAC 2452
Db 1621 ATGCCGAAGCGGCGCGGTTCAAGGGGTGCGCGCGCATCAATCTATCGAACGCTCTCAAC 1680
Qy 2453 CAGGTGCTCCGCTGGCGGCTGGGTCCGTCGAGATCTTCTTCAGCGGCAAGGCCCGCTG 2512
Db 1681 CAGGTGCTCCGCTGGCGGCTGGGTCCGTCGAGATCTTCTTCAGCGGCAAGGCCCGCTG 1740
Qy 2513 CTGTACGGCTACAGAAACGGCAACCTCAAGTGGCTGGAGCGTTCGCTTACATCAACACC 2572
Db 1741 CTGTACGGCTACAGAAACGGCAACCTCAAGTGGCTGGAGCGTTCGCTTACATCAACACC 1800
Qy 2573 ACCATCTACCTTCACTCTCGCTCCGCTGTCTGCTACTGCAACCTCCCGCGCTGTGC 2632
Db 1801 ACCATCTACCTTCACTCTCGCTCCGCTGTCTGCTACTGCAACCTCCCGCGCTGTGC 1860
Qy 2633 CTCTCACCGGCAAGTTCACTCATGCTGCTAGTACGCTTCGCCAGCTCTTCTTCATC 2692
Db 1861 CTCTCACCGGCAAGTTCACTCATGCTGCTAGTACGCTTCGCCAGCTCTTCTTCATC 1920
Qy 2693 GCGCTCTTCATGTCATCTTCGCGAGCGGCATCTCGAGATCGGTGGAGCGGGGTGAGC 2752
Db 1921 GCGCTCTTCATGTCATCTTCGCGAGCGGCATCTCGAGATCGGTGGAGCGGGGTGAGC 1980
Qy 2753 ATCGAGGAGTGTGGAGGAAACGAGCAGTTCTGGGTCTATCGCGCGGTGTCCGGCATCTC 2812
Db 1981 ATCGAGGAGTGTGGAGGAAACGAGCAGTTCTGGGTCTATCGCGCGGTGTCCGGCATCTC 2040
Qy 2813 TTGCGCGCTGTCAGGGCTGTCTCAAGTCTTCGCGGGATCGACACCAACTTCAAGTGC 2872
Db 2041 TTGCGCGCTGTCAGGGCTGTCTCAAGTCTTCGCGGGATCGACACCAACTTCAAGTGC 2100
Qy 2873 ACTTCAAGGCCACCGGGACGAGACGAGATTCGCGAGTCTTACGCTTCAAGTGC 2932
Db 2101 ACTTCAAGGCCACCGGGACGAGACGAGATTCGCGAGTCTTACGCTTCAAGTGC 2160
Qy 2933 ACCAGCTCTCTCATCCGCGCCACCAAGCTGTCTCATCATTAACGCTCATCGCGCTGTG 2992
Db 2161 ACCAGCTCTCTCATCCGCGCCACCAAGCTGTCTCATCATTAACGCTCATCGCGCTGTG 2220
Qy 2993 GGCATCTCGAGCGCCATCAACAGGCTACGAGTCTTGGGGCGCTCTTCGGCAGCTC 3052
Db 2221 GGCATCTCGAGCGCCATCAACAGGCTACGAGTCTTGGGGCGCTCTTCGGCAGCTC 2280
Qy 3053 TTCTTCGCTCTGCGTCTATCGTCAACCTCTACCGCTTCTCAAGGGGCTCATTGGGCGC 3112
Db 2281 TTCTTCGCTCTGCGTCTATCGTCAACCTCTACCGCTTCTCAAGGGGCTCATTGGGCGC 2340
Qy 3113 CAGAACAGGACCGCCACCGCTTGTGTCTCATCTGGTCAATCTGCTGGCTCTCATCTTCTCC 3172
Db 2341 CAGAACAGGACCGCCACCGCTTGTGTCTCATCTGGTCAATCTGCTGGCTCTCATCTTCTCC 2400
Qy 3173 CTGCTCTGGGTGAGATCGACCTTTCATCTGTCAGGACCAAGGCCCGGACGTCAGGACG 3232
Db 2401 CTGCTCTGGGTGAGATCGACCTTTCATCTGTCAGGACCAAGGCCCGGACGTCAGGACG 2460
Qy 3233 TGTGGCATCAATTTGCTGAGCTGTATTAAAGTTTCAAAATTTCTGAGCTTTGTGCATAGGG 3292
Db 2461 TGTGGCATCAATTTGCTGAGCTGTATTAAAGTTTCAAAATTTCTGAGCTTTGTGCATAGGG 2520
Qy 3293 AGAAAAAACAATTTAGAAATTTGTAGGTTTGTGTGTCTGTAAATGTTATGATGCCAG 3352
Db 2521 AGAAAAAACAATTTAGAAATTTGTAGGTTTGTGTGTCTGTAAATGTTATGATGCCAG 2580
Qy 3353 AATTGTGCGACGAGGAAATTGAACAAAGGACAAAGGTTTGAATTTAAATGGCAAA 3406
Db 2581 AATTGTGCGACGAGGAAATTGAACAAAGGACAAAGGTTTGAATTTAAATGGCAAA 2634


```
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 139789
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_58974C.1
US-10-425-115-139789

Query Match      5.0%; Score 173; DB 18; Length 2830;
Best Local Similarity 98.9%; Pred. No. 7.7e-77;
Matches 373; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1457 ATCAACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1516
Db      |||||||
Qy 826 ATCAACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 885
Db      |||||||
Qy 1517 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1576
Db      |||||||
Qy 886 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 945
Db      |||||||
Qy 1577 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1636
Db      |||||||
Qy 946 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1005
Db      |||||||
Qy 1637 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1696
Db      |||||||
Qy 1006 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1065
Db      |||||||
Qy 1697 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1756
Db      |||||||
Qy 1066 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1125
Db      |||||||
Qy 1757 ATCAACACAGCAAGGCGCATCCGGAGGCGCATGCTTCTCATGAGCCCTCAGGTCGGC 1816
Db      |||||||
Qy 1126 ATCAACACAGCAAGGCGCATCCGGAGGCGCATGCTTCTCATGAGCCCTCAGGTCGGC 1185
Db      |||||||
Qy 1817 CGGAAGGTCTGCTACGT 1833
Db      |||||||
Qy 1186 CGGAAGGTCTGCTACGT 1202
Db      |||||||

RESULT 5
US-10-425-114-24930
; Sequence 24930, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53131)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24930
; LENGTH: 3589
; TYPE: DNA
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-221-C10_FLI
US-10-425-114-24930

Query Match      5.0%; Score 173; DB 17; Length 3589;
Best Local Similarity 98.9%; Pred. No. 7.5e-77;
Matches 373; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1457 ATCAACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1516
Db      |||||||
Qy 1615 ATCAACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1674
Db      |||||||
Qy 1517 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1576
Db      |||||||
Qy 1675 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1734
Db      |||||||
Qy 1577 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1636
Db      |||||||
Qy 1735 GGCACGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1794
Db      |||||||
Qy 1637 CGTGAGAGCGCGCGGATTCAGACACCAAGAGGCGCGGCGCATGAACGCTCTGATT 1696
Db      |||||||
Qy 1795 CGTGAGAGCGCGCGGATTCAGACACCAAGAGGCGCGGCGCATGAACGCTCTGATT 1854
Db      |||||||
Qy 1697 CGCGTCTCCGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1756
Db      |||||||
Qy 1855 CGCGTCTCCGCGCTGGTGGCCAAAGCCATGAAGTGGCGGAGGGGTGGATCATGAAGAC 1914
Db      |||||||
Qy 1757 ATCAACACAGCAAGGCGCATCCGGAGGCGCATGCTTCTCATGAGCCCTCAGGTCGGC 1816
Db      |||||||
Qy 1915 ATCAACACAGCAAGGCGCATCCGGAGGCGCATGCTTCTCATGAGCCCTCAGGTCGGC 1974
Db      |||||||
Qy 1817 CGGAAGGTCTGCTACGT 1833
Db      |||||||
Qy 1975 CGGAAGGTCTGCTACGT 1991
Db      |||||||

RESULT 6
US-10-425-115-179459
; Sequence 179459, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 179459
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(335)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION:
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9524C.1
US-10-425-115-179459

Query Match      3.6%; Score 124; DB 18; Length 335;
Best Local Similarity 100.0%; Pred. No. 6.7e-52;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCGTCGCGCTGCTCGCAATCGCAATCTGTCGAGCACCTGAGGGGTGGAGGCCGAG 60
Db      |||||||
Qy 126 CTGCGTCGCGCTGCTCGCAATCGCAATCTGTCGAGCACCTGAGGGGTGGAGGCCGAG 185
Db      |||||||
```

```
QY 61 AGCTAGCTAGCAGCGCGCTCGCGCGCGATGAGAGCCAGCGCGCGGTGCTGCGCGG 120
Db 186 AGTTCAGCTAGCAGCGCGCTCGCGCGCGATGAGAGCCAGCGCGCGGTGCTGCGCGG 245
QY 121 CTCG 124
Db 246 CTCG 249

RESULT 7
US-10-767-701-10062
; Sequence 10062, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 10062
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; NAME/KEY: unsure
; LOCATION: (1)..(999)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS50500_1
US-10-767-701-10062

Query Match 3.4%; Score 116; DB 18; Length 999;
Best Local Similarity 100.0%; Pred. No. 6.8e-48;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2312 GCGTGGATCTACGGGTGATCAGGAGGACATCTGACGGGGTTCAAGATGCACTGCCG 2371
Db 35 GCGTGGATCTACGGGTGATCAGGAGGACATCTGACGGGGTTCAAGATGCACTGCCG 94

QY 2372 GCGTGGCGCTCCGTGACTGCATCCGAAAGCGCGCGCTTCAAGGGTCCGCGCC 2427
Db 95 GCGTGGCGCTCCGTGACTGCATCCGAAAGCGCGCGCTTCAAGGGTCCGCGCC 150

RESULT 8
US-10-425-115-10141
; Sequence 10141, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 10141
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: unsure
; LOCATION: (1)..(633)
; OTHER INFORMATION: Clone ID: MRT4577_10924C.1
US-10-425-115-10141

Query Match 3.2%; Score 110; DB 18; Length 633;
Best Local Similarity 99.4%; Pred. No. 7.9e-45;
```

```
Matches 160; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 281 TGC CGG C C C T G C T A C G A G T A C G A G C C C G G G A G G G C A C C A G A A C T G C C C C A G T C C A A G 340
Db 436 TGC CGG C C C T G C T A C G A G T A C G A G C C C G G G A G G G C A C C A G A A C T G C C C C A G T C C A A G 495
QY 341 A C G C G C T A C A A G C G C C T C A A G G G A G C C C A G G G T T G C C G G G A C G A T A C C A G G A G G A C 400
Db 496 A C G C G C T A C A A G C G C C T C A A G G G A G C C C A G G G T T G C C G G G A C G A T A C C A G G A G G A C 555
QY 401 A T C G A C G A C C T G G A C A C G A G T T C A A C A T C G A C G A C G A A 441
Db 556 A T C G A C G A C C T G G A C A C G A G T T C A A C A T C G A C G A C G A A 596

RESULT 9
US-10-437-963-15093/c
; Sequence 15093, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15093
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20969C.1
US-10-437-963-15093

Query Match 2.4%; Score 83; DB 18; Length 3448;
Best Local Similarity 100.0%; Pred. No. 3.3e-31;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 813 C A C T G A A C C A C G A G G C G A G C C G C T G T C G A G G A A G G T G C G A T C G C G T C G A G C A A G G 872
Db 2646 C A C T G A A C C A C G A G G C G A G C C G C T G T C G A G G A A G G T G C G A T C G C G T C G A G C A A G G 2587

QY 873 T G A A C C C G T A C C G G A T G G T G A T C 895
Db 2586 T G A A C C C G T A C C G G A T G G T G A T C 2564

RESULT 10
US-10-767-701-29317
; Sequence 29317, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 29317
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Sorghum bicolor
```

FEATURE:
OTHER INFORMATION: Clone ID: 9298823
US-10-767-701-29317

Query Match 1.7%; Score 59; DB 18; Length 549;
Best Local Similarity 100.0%; Pred. No. 5.8e-19; Indels 0; Gaps 0;
Matches 59; Conservative 0; Mismatches 0;

Qy 1733 CTCACCTGGACTGTGATCACTACATCAACACGAGGCGCATCCGGAGGCCATGTG 1791
Db 1 CTCACCTGGACTGTGATCACTACATCAACACGAGGCGCATCCGGAGGCCATGTG 59

RESULT 11

US-10-425-115-139788
Sequence 139788, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 139788
LENGTH: 488
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_58973C.1
US-10-425-115-139788

Query Match 1.7%; Score 58; DB 18; Length 488;
Best Local Similarity 100.0%; Pred. No. 1.9e-18; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0;

Qy 1077 CCCTCAGTACGAGAGGAGGCGCGTCTGCTGCGGCGGTGGACCTGTTCGT 1134
Db 201 CCCTCAGTACGAGAGGAGGCGCGTCTGCTGCGGCGGTGGACCTGTTCGT 258

RESULT 12

US-09-900-237-5
Sequence 5, Application US/09900237
Patent No. US20020120124A1
GENERAL INFORMATION:
APPLICANT: Allen, Stephen
TITLE OF INVENTION: Plant Cellulose Synthases
FILE REFERENCE: Bb1170 US CIP
CURRENT APPLICATION NUMBER: US/09/900,237
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: 60/092,844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US99/15871
PRIOR FILING DATE: 1999-07-13
PRIOR APPLICATION NUMBER: 09/720383
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Microsoft Office 97
SEQ ID NO 5
LENGTH: 1189
TYPE: DNA
ORGANISM: Zea mays
US-09-900-237-5

Query Match 1.5%; Score 50; DB 9; Length 1189;
Best Local Similarity 100.0%; Pred. No. 2e-14; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0;

Qy 2771 AACGAGCAGTTCTGGGTCAATCGCGGGGTGTCCGGCATCTCTTCGCCGT 2820
Db 485 AACGAGCAGTTCTGGGTCAATCGCGGGGTGTCCGGCATCTCTTCGCCGT 534

RESULT 13

US-10-425-114-2510
Sequence 2510, Application US/10425114
Publication No. US20040034889A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 2510
LENGTH: 2351
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700214676_FLI
US-10-425-114-2510

Query Match 1.5%; Score 50; DB 17; Length 2351;
Best Local Similarity 100.0%; Pred. No. 1.9e-14; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0;

Qy 2771 AACGAGCAGTTCTGGGTCAATCGCGGGGTGTCCGGCATCTCTTCGCCGT 2820
Db 1674 AACGAGCAGTTCTGGGTCAATCGCGGGGTGTCCGGCATCTCTTCGCCGT 1723

RESULT 14

US-10-425-115-137279
Sequence 137279, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 137279
LENGTH: 2684
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(2684)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_56676C.1
US-10-425-115-137279

Query Match 1.5%; Score 50; DB 18; Length 2684;
Best Local Similarity 100.0%; Pred. No. 1.8e-14; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0;

Qy 2771 AACGAGCAGTTCTGGGTCAATCGCGGGGTGTCCGGCATCTCTTCGCCGT 2820
Db 1843 AACGAGCAGTTCTGGGTCAATCGCGGGGTGTCCGGCATCTCTTCGCCGT 1892

```
RESULT 15
US-10-209-059-25
; Sequence 25, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3470
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-25

Query Match          1.5%; Score 50; DB 16; Length 3470;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2771 AACGAGCAGTTCGGGTTCATCGCGCGGTGTCGCGCATCTCTTCGCCGT 2820
    |||||||
Db 2786 AACGAGCAGTTCGGGTTCATCGCGCGGTGTCGCGCATCTCTTCGCCGT 2835

RESULT 16
US-10-627-132-25
; Sequence 25, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3470
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-25

Query Match          1.5%; Score 50; DB 17; Length 3470;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2771 AACGAGCAGTTCGGGTTCATCGCGCGGTGTCGCGCATCTCTTCGCCGT 2820
    |||||||
Db 2786 AACGAGCAGTTCGGGTTCATCGCGCGGTGTCGCGCATCTCTTCGCCGT 2835
```

```
RESULT 17
US-09-955-999-23
; Sequence 23, Application US/09955599
; Publication No. US20030036505A1
; GENERAL INFORMATION:
; APPLICANT: Barash et al.
; TITLE OF INVENTION: Signal Transduction Pathway Component Polynucleotides, Polypeptides, Antibodies, and Methods Based Thereon
; FILE REFERENCE: PT086P1
; CURRENT APPLICATION NUMBER: US/09/955,999
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (625)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (787)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-955-999-23

Query Match          1.4%; Score 48; DB 10; Length 907;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
    |||||||
Db 859 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 906

RESULT 18
US-10-437-963-39652
; Sequence 39652, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 39652
; LENGTH: 1657
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43170C.1
US-10-437-963-39652

Query Match          1.4%; Score 48; DB 18; Length 1657;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3396 TAAATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
```


; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 378
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1799)..(1799)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-378

Query Match 1.4%; Score 47; DB 15; Length 1826;
Best Local Similarity 100.0%; Pred. No. 6.3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3397 AATGGCAA 3443
|||||
Db 1740 AATGGCAA 1786

RESULT 24
US-10-425-115-55122/c
; Sequence 55122, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 55122
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_150269C.1
US-10-425-115-55122

Query Match 1.3%; Score 46; DB 18; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3398 AATGGCAA 3443
|||||
Db 84 AATGGCAA 39

RESULT 25
US-09-925-298-401
; Sequence 401, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 401
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (162)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (165)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (187)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-298-401

Query Match 1.3%; Score 45; DB 9; Length 189;
Best Local Similarity 100.0%; Pred. No. 8.1e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3399 ATGGCAA 3443
|||||
Db 104 ATGGCAA 148

RESULT 26
US-10-102-806-401
; Sequence 401, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 401
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (162)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (165)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature


```
; LOCATION: (166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (187)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-401

Query Match
Best Local Similarity 1.3%; Score 45; DB 14; Length 189;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 104 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 148

RESULT 27
US-10-425-115-102183
; Sequence 102183, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 102183
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(358)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_24699C.1
US-10-425-115-102183

Query Match
Best Local Similarity 1.3%; Score 45; DB 18; Length 358;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 91 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 135

RESULT 28
US-09-918-995-11420
; Sequence 11420, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11420
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1)...(480)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11420

Query Match
Best Local Similarity 1.3%; Score 45; DB 10; Length 480;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 208 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 252

RESULT 29
US-10-425-115-52426
; Sequence 52426, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 52426
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_147812C.1
US-10-425-115-52426

Query Match
Best Local Similarity 1.3%; Score 45; DB 18; Length 711;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 77 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 121

RESULT 30
US-10-425-115-105831
; Sequence 105831, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 105831
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(864)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_28018C.1
US-10-425-115-105831

Query Match
Best Local Similarity 1.3%; Score 45; DB 18; Length 864;
```

Best Local Similarity 100.0%; Pred. No. 7e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3399 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3443
Db 775 ATGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 819

Search completed: March 12, 2005, 10:42:43
Job time : 1226 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2005, 02:20:28 ; Search time 7010 Seconds
(without alignments)
18695.475 Million cell updates/sec

Title: US-10-627-132-29

Perfect score: 3443

Sequence: 1 ctgcgtcgccctgcctgcga.....aaaaaaaaaaaaaaaaaaaa 3443

Scoring table: OLIGO_NUC

Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 25

Total number of hits satisfying chosen parameters: 596789

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_est4.*

5: gb_est5.*

6: gb_est6.*

7: gb_ges1.*

8: gb_ges2.*

9: gb_ges2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	745	21.6	943	9	CG454544
C 2	694	20.2	872	8	CG427726
C 3	643	18.7	949	9	CG211396
C 4	583	16.9	822	9	CG198455
C 5	368	10.7	480	2	BE050991
C 6	356	10.3	631	8	CG427725
C 7	356	10.3	739	9	CG145359
C 8	356	10.3	782	9	CG371975
C 9	356	10.3	828	9	CG619840
C 10	356	10.3	835	9	CG300337
C 11	356	10.3	867	9	CG081564
C 12	356	10.3	917	9	CG263400
C 13	356	10.3	945	9	CG300346
C 14	349	10.1	818	8	CG446974
C 15	332	9.6	452	9	CG656950
C 16	320	9.3	608	6	CA183623
C 17	320	9.3	624	6	CA173049
C 18	318	9.2	808	9	CG436514
C 19	305	8.9	748	9	CG454501
C 20	269	7.8	640	9	CG705263
C 21	269	7.8	833	8	CG446976
C 22	269	7.8	841	8	BZ733711
C 23	266	7.7	651	8	CC389096
C 24	266	7.7	815	9	CC705271

CC655942	818	7.7	266	C 25	CG655942
CG343883	7.7	266	266	C 26	CG343883
CG145361	7.6	261	261	C 27	CG145361
BE050992	7.1	246	246	C 28	BE050992
BZ733703	6.9	236	236	C 29	BZ733703
CG371985	6.2	214	214	C 30	CG371985
CG430615	6.2	214	214	C 31	CG430615
CG263391	6.2	214	214	C 32	CG263391
CG343868	6.2	214	214	C 33	CG343868
CG434216	6.2	214	214	C 34	CG434216
CG434218	6.2	214	214	C 35	CG434218
CG685594	6.0	206	206	C 36	CG685594
CG685588	6.0	206	206	C 37	CG685588
CG619834	6.0	206	206	C 38	CG619834
CG081561	6.0	206	206	C 39	CG081561
CC000780	5.9	203	203	C 40	CC000780
CA173106	5.8	200	200	C 41	CA173106
CC374709	5.3	183	183	C 42	CC374709
CA194818	5.3	182	182	C 43	CA194818
CG712515	5.2	178	178	C 44	CG712515
CG081196	5.0	173	173	C 45	CG081196
BZ640404	5.0	173	173	C 46	BZ640404
CC651626	5.0	173	173	C 47	CC651626
CG615542	5.0	173	173	C 48	CG615542
CG370129	5.0	173	173	C 49	CG370129
CF433508	4.9	167	167	C 50	CF433508
CA165585	4.8	158	158	C 51	CA165585
CA155363	4.6	158	158	C 52	CA155363
CA261021	4.3	149	149	C 53	CA261021
BZ640411	4.3	148	148	C 54	BZ640411
CG381495	4.2	146	146	C 55	CG381495
CG660677	4.2	146	146	C 56	CG660677
CG381504	4.2	146	146	C 57	CG381504
CA133799	4.2	146	146	C 58	CA133799
CO519653	3.9	134	134	C 59	CO519653
CG374710	3.8	134	134	C 60	CG374710
CA235459	3.7	126	126	C 61	CA235459
CG712518	3.6	125	125	C 62	CG712518
CA127439	3.5	122	122	C 63	CA127439
CA208554	3.5	122	122	C 64	CA208554
CA129877	3.5	122	122	C 65	CA129877
CF650564	3.4	118	118	C 66	CF650564
BE357401	3.4	116	116	C 67	BE357401
CF430024	3.4	116	116	C 68	CF430024
CA268135	3.3	115	115	C 69	CA268135
BZ586599	3.3	113	113	C 70	BZ586599
CL161384	3.3	113	113	C 71	CL161384
CL222797	3.1	108	108	C 72	CL222797
BI141114	3.1	107	107	C 73	BI141114
BE357339	3.1	107	107	C 74	BE357339
CG081198	3.1	107	107	C 75	CG081198
CL167974	3.1	107	107	C 76	CL167974
CA253951	3.0	102	102	C 77	CA253951
CG359736	2.9	101	101	C 78	CG359736
CG038978	2.9	101	101	C 79	CG038978
AQ844522	2.8	98	98	C 80	AQ844522
BE125071	2.7	93	93	C 81	BE125071
CA172230	2.7	92	92	C 82	CA172230
CA244716	2.7	92	92	C 83	CA244716
CA069753	2.7	92	92	C 84	CA069753
CC722849	2.7	92	92	C 85	CC722849
CL198038	2.7	92	92	C 86	CL198038
CG122858	2.7	92	92	C 87	CG122858
CG615547	2.7	92	92	C 88	CG615547
CC704091	2.5	85	85	C 89	CC704091
CC350870	2.5	85	85	C 90	CC350870
CO519652	2.4	84	84	C 91	CO519652
CA069829	2.4	83	83	C 92	CA069829
CB646549	2.4	83	83	C 93	CB646549
CL976367	2.4	83	83	C 94	CL976367
CG660675	2.2	77	77	C 95	CG660675
CA185787	2.1	74	74	C 96	CA185787
BU255931	2.1	74	74	C 97	BU255931

98 72 2.1 981 9 CG370118 CG370118 OGVC052TH
99 71 2.1 409 9 CL161383 CL161383 104_352.1
100 71 2.1 429 6 CA291358 CA291358 SCCCN200

ALIGNMENTS

RESULT 1
CG454544/c 943 bp DNA linear GSS 17-SEP-2003
LOCUS PUIKA96TDB ZM_0.6_1.0_KB Zea mays genomic clone ZMBTa0601023,
DEFINITION genomic survey sequence.

ACCESSION CG454544
VERSION CG454544.1 GI:34839544
KEYWORDS GSS.

SOURCE

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 943)

Whitealaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and

Bennetzen,J.

Maize Genomics Consortium

Unpublished (2003)

Other GSSs: PUIKA96TDB PUIKA96TBB

Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1. .943

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone_lib="ZMBTa0601023"

/clone_lib="ZM_0.6_1.0_KB"

/notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high

Cot selected genomic DNA library"

ORIGIN

Query Match 21.6%; Score 745; DB 9; Length 943;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 745; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2662 GATTACAGGTTCCGAGCCTCTTCTTCATCGCCCTTTCATCTCCATCTTCGCGAGGG 2721
|||||

DB 903 GATTACAGGTTCCGAGCCTCTTCTTCATCGCCCTTTCATCTCCATCTTCGCGAGGG 844
|||||

QY 2722 CATCTCTGGAGATCGGTGGAGCGGGTGAGCATCGAGGAGTGGTGGAGGAACGAGCAGTT 2781
|||||

DB 843 CATCTCTGGAGATCGGTGGAGCGGGTGAGCATCGAGGAGTGGTGGAGGAACGAGCAGTT 784
|||||

QY 2782 CTGGGTCATCGCGCGGTGTCGCGCATCTCTTCGCGGTGGTGCAGGGCTGCTCAAGTT 2841
|||||

DB 783 CTGGGTCATCGCGCGGTGTCGCGCATCTCTTCGCGGTGGTGCAGGGCTGCTCAAGTT 724
|||||

QY 2842 CCTCGCGGGATCGACCAACTTCCACGTCACCTCCAGGCCACCGGCGAGGAGCA 2901
|||||

DB 723 CCTCGCGGGATCGACCAACTTCCACGTCACCTCCAGGCCACCGGCGAGGAGCA 664
|||||

QY 2902 CGAGTTTCGCGAGCTCTACGCTTCAAGTGGACCGCTCTCATCCCGCCACACGCT 2961
|||||

DB 663 CGAGTTTCGCGAGCTCTACGCTTCAAGTGGACCGCTCTCATCCGCGCCACGCT 604
|||||

QY 2962 GCTCATCTTAACGTCATCGGCGTGTGGCCGCGCATCTCGGACGCCATCAACACGGGTA 3021
|||||

DB 603 GCTCATCTTAACGTCATCGGCGTGTGGCCGCGCATCTCGGACGCCATCAACACGGGTA 544
|||||

QY 3022 CAGTCCCTGGGGGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTTCATCGTCCACCT 3081
|||||

DB 543 CAGTCCCTGGGGGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTTCATCGTCCACCT 484
|||||

QY 3082 CTACCCGTTTCCTCAAGGGGCTCATGGGGGCCAGAACAGAGCGCCACCGTTGTTTCAT 3141
|||||

DB 483 CTACCCGTTTCCTCAAGGGGCTCATGGGGGCCAGAACAGAGCGCCACCGTTGTTTCAT 424
|||||

QY 3142 CTGGTCCATCTCTGGGCTCTCATCTTCTCCCTGCTCTGGGTACAGATCGACCCCTTTCAT 3201
|||||

DB 423 CTGGTCCATCTCTGGGCTCTCATCTTCTCCCTGCTCTGGGTACAGATCGACCCCTTTCAT 364
|||||

QY 3202 CGTCAGGACCAAGGGGCCGAGCTCAGGCAAGTGGCATCAATTCGTGAGCTGTTTATTA 3261
|||||

DB 363 CGTCAGGACCAAGGGGCCGAGCTCAGGCAAGTGGCATCAATTCGTGAGCTGTTTATTA 304
|||||

QY 3262 AGGTTCAAAATTCGAGCTTGTGCATAGGGAGAAAAAACAATTTAGAAAATTTCTAAG 3321
|||||

DB 303 AGGTTCAAAATTCGAGCTTGTGCATAGGGAGAAAAAACAATTTAGAAAATTTCTAAG 244
|||||

QY 3322 GTTGTGCTCTGTAATGTTATGTTACCCAGAAATTCGAGCGAGGAATTGAACAAGGA 3381
|||||

DB 243 GTTGTGCTCTGTAATGTTATGTTACCCAGAAATTCGAGCGAGGAATTGAACAAGGA 184
|||||

QY 3382 CAAAGTTTGATTTAAATGGCAAA 3406
|||||

DB 183 CAAAGTTTGATTTAAATGGCAAA 159
|||||

RESULT 2
CC427726/c 872 bp DNA linear GSS 19-MAY-2003
LOCUS PUHAM29TD ZM_0.6_1.0_KB Zea mays genomic clone ZMBTa407E09,
DEFINITION genomic survey sequence.

ACCESSION CC427726

VERSION CC427726.1 GI:30907816

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 872)

Whitealaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and

Bennetzen,J.

Maize Genomics Consortium

Unpublished (2003)

Other GSSs: PUHAM29TB

Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1. .872

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone_lib="ZMBTa407E09"

/clone_lib="ZM_0.6_1.0_KB"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Query Match 20.2%; Score 694; DB 8; Length 872;


```
Db      251 TCAGTTCGCGAGAGTTTCGACGGCATCGAGTGCACGACGATACGTAACAGGAACAC 310
Qy      1894 CGTCTTCTTGACATCAACATGAAGGGCTGGAAGGCATCAAGGCCCGGGTGTAGCTCGG 1953
Db      311 CGTCTTCTTGACATCAACATGAAGGGCTGGAAGGCATCAAGGCCCGGGTGTAGCTCGG 370
Qy      1954 GACAGGGTGCCTGTTCCGGCGCAGGGCTCTACGGCTACAACTCCCAAGGAGACCAA 2013
Db      371 GACAGGGTGCCTGTTCCGGCGCAGGGCTCTACGGCTACAACTCCCAAGGAGACCAA 430
Qy      2014 GAGGCCCAAGATGCTGACCTCGCACTGCTGCCGCTGCTTCCGCCGCAAGA 2063
Db      431 GAGGCCCAAGATGCTGACCTCGCACTGCTGCCGCTGCTTCCGCCGCAAGA 480

RESULT 6
LOCUS   CC427725
DEFINITION PUHAM29TB ZM 0.6_1.0 KB Zea mays genomic clone ZMMBTA407E09,
          genomic survey sequence.
ACCESSION CC427725
VERSION   CC427725.1 GI:30907815
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 631)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
          Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Other GSSs: PUHAM29TD
          Contact: Cathy Whitelaw
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TR
          Class: sheared ends.
FEATURES  Location/Qualifiers
            source          1..631
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /strain="B73"
                        /db_xref="taxon:4577"
                        /clone_lib="ZMMBTA407E09"
                        /clone="ZM 0.6_1.0 KB"
                        /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                        CoT selected genomic DNA library"

ORIGIN
Query Match      10.3%; Score 356; DB 8; Length 631;
Best Local Similarity 100.0%; Pred. No. 6.5e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2307 AGCTGGGGTGGATCTACGGTTCGATCAGGAGGACATCCTGACGGGGTTCAAGATGCAC 2366
Db      54  AGCTGGGGTGGATCTACGGTTCGATCAGGAGGACATCCTGACGGGGTTCAAGATGCAC 113
Qy      2367 GCCCGGGTGGCGCTCCGCTGCTACTGTCATGCCAAGCGGGCGGCTTCAAGGGGTGGCGC 2426
Db      114 GCCCGGGTGGCGCTCCGCTGCTACTGTCATGCCAAGCGGGCGGCTTCAAGGGGTGGCGC 173
Qy      2427 CGATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTCGGTTCGAGA 2486
Db      174 CGATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTCGGTTCGAGA 233
Qy      2487 TCCTTTCACGGCGCACAGCCCCCTGCTGTAGCGGTACAGGACCGGCAACCTCAAGTGGC 2546
```

```
Db      234 TCCTTTTCAGCCCGCACAGCCCCCTGCTGTACGGCTACAAGAACGGCAACCTCAAGTGGC 293
Qy      2547 TGGAGCGCTTTCGCTTACATCAACACCACTATACCCCTTACCTCGCTCCCGCTGCTCG 2606
Db      294 TGGAGCGCTTTCGCTTACATCAACACCACTATACCCCTTACCTCGCTCCCGCTGCTCG 353
Qy      2607 CCTACTGCACCTTCCTCCCGCGCTGTCTCTCTACCGGCAAGTTTCATCATGCGCTCG 2662
Db      354 CCTACTGCACCTTCCTCCCGCGCTGTCTCTCTACCGGCAAGTTTCATCATGCGCTCG 409

RESULT 7
LOCUS   CG145359
DEFINITION PUJEW43TB ZM 0.6_1.0 KB Zea mays genomic clone ZMMBTA0652H13,
          genomic survey sequence.
ACCESSION CG145359
VERSION   CG145359.1 GI:34036142
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 739)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
          Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Other GSSs: PUJEW43TD
          Contact: Cathy Whitelaw
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TR
          Class: sheared ends.
FEATURES  Location/Qualifiers
            source          1..739
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /strain="B73"
                        /db_xref="taxon:4577"
                        /clone_lib="ZMMBTA0652H13"
                        /clone="ZM 0.6_1.0 KB"
                        /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                        CoT selected genomic DNA library"

ORIGIN
Query Match      10.3%; Score 356; DB 9; Length 739;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2307 AGCTGGGGTGGATCTACGGTTCGATCAGGAGGACATCCTGACGGGGTTCAAGATGCAC 2366
Db      112 AGCTGGGGTGGATCTACGGTTCGATCAGGAGGACATCCTGACGGGGTTCAAGATGCAC 171
Qy      2367 GCCCGGGTGGCGCTCCGCTGCTACTGTCATGCCAAGCGGGCGGCTTCAAGGGGTGGCGC 2426
Db      172 GCCCGGGTGGCGCTCCGCTGCTACTGTCATGCCAAGCGGGCGGCTTCAAGGGGTGGCGC 231
Qy      2427 CGATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTCGGTTCGAGA 2486
Db      232 CGATCAATCTATCGGACCGTCTCAACAGAGTGTCTCCGGTGGCGCTCGGTTCGAGA 291
Qy      2487 TCCTTTCACGGCGCACAGCCCCCTGCTGTAGCGGTACAGGACCGGCAACCTCAAGTGGC 2546
Db      292 TCCTTTCACGGCGCACAGCCCCCTGCTGTAGCGGTACAGGACCGGCAACCTCAAGTGGC 351
Qy      2547 TGGAGCGCTTTCGCTTACATCAACACCACTATACCCCTTCACTCGCTCCCGCTGCTCG 2606
```

```
||||| 352 TGGAGCGCTTGGCTTACATCAACACCACTTACCCCTTACCTCGCTCCGCTGCTCG 411
|||||
QY 2607 CCTACTGCACCTCCCGCGCTGCTCTCTACCGGCAAGTTTCATCATGCGGTGCG 2662
|||||
Db 412 CCTACTGCACCTCCCGCGCTGCTCTCTACCGGCAAGTTTCATCATGCGGTGCG 467
|||||

RESULT 8
CG371975/c
LOCUS CG371975 782 bp DNA linear GSS 26-AUG-2003
DEFINITION OG1AD62TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBma0715L04,
genomic survey sequence.
ACCESSION CG371975
VERSION CG371975.1 GI:34289242
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 782)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OG1AD62TV
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..782
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0715L04"
/clone_lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 782;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2307 AGCTGGGGTGGATCTACGGGTGCATCAGGAGACATCTCTGACGGGTTCAAGATGCAC 2366
|||||
Db 609 AGCTGGGGTGGATCTACGGGTGCATCAGGAGACATCTCTGACGGGTTCAAGATGCAC 550
|||||
QY 2367 GCCCGGGTGGCGCTCGGTGTACTGCAATCCGAAGCGGGCGGTTCAGAGGGTCCGGCGC 2426
|||||
Db 549 GCCCGGGTGGCGCTCGGTGTACTGCAATCCGAAGCGGGCGGTTCAGAGGGTCCGGCGC 490
|||||
QY 2427 CGATCAATCTATCGGACCGGTCTCAACAGGTGTCTCCGGTGGGCGCTCGGTTCGAGA 2486
|||||
Db 489 CGATCAATCTATCGGACCGGTCTCAACAGGTGTCTCCGGTGGGCGCTCGGTTCGAGA 430
|||||
QY 2487 TCTTCTTACGGCGCAGCGCCCTGCTGTACCGGTACAGAGACGGCAACCTCAAGTGGC 2546
|||||
Db 429 TCTTCTTACGGCGCAGCGCCCTGCTGTACCGGTACAGAGACGGCAACCTCAAGTGGC 370
|||||
QY 2547 TGGAGCGCTTCGGCTATACCAACCACTTACCCCTTACCTCGCTCCCGCTGCTCG 2606
|||||
Db 369 TGGAGCGCTTCGGCTATACCAACCACTTACCCCTTACCTCGCTCCCGCTGCTCG 310
|||||
QY 2607 CCTACTGCACCTCCCGCGCTGTGCTCTCTCAACAGGTTCATCATGCGGTGCG 2662
|||||
```

```
||||| 309 CCTACTGCACCTCCCGCGCTGCTCTCTCACCGCAAGTTTCATCATGCGGTGCG 254
|||||

RESULT 9
CG619840/c
LOCUS CG619840 828 bp DNA linear GSS 19-JUN-2003
DEFINITION OGULU56TV ZM 0.7_1.5_KB Zea mays genomic clone ZMMBma0470J15,
genomic survey sequence.
ACCESSION CG619840
VERSION CG619840.1 GI:31986261
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 828)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGULU56TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: RF
Class: sheared ends.
FEATURES
source
1..828
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0470J15"
/clone_lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 828;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2307 AGCTGGGGTGGATCTACGGGTGCATCAGGAGACATCTCTGACGGGTTCAAGATGCAC 2366
|||||
Db 453 AGCTGGGGTGGATCTACGGGTGCATCAGGAGACATCTCTGACGGGTTCAAGATGCAC 394
|||||
QY 2367 GCCCGGGTGGCGCTCCGTGTACTGCAATCCGAAGCGGGCGGTTCAGAGGGTCCGGCGC 2426
|||||
Db 393 GCCCGGGTGGCGCTCCGTGTACTGCAATCCGAAGCGGGCGGTTCAGAGGGTCCGGCGC 334
|||||
QY 2427 CGATCAATCTATCGGACCGGTCTCAACAGGTGTCTCCGGTGGGCGCTCGGTTCGAGA 2486
|||||
Db 333 CGATCAATCTATCGGACCGGTCTCAACAGGTGTCTCCGGTGGGCGCTCGGTTCGAGA 274
|||||
QY 2487 TCTTCTTACGGCGCAGCGCCCTGCTGTACCGGTACAGAGACGGCAACCTCAAGTGGC 2546
|||||
Db 273 TCTTCTTACGGCGCAGCGCCCTGCTGTACCGGTACAGAGACGGCAACCTCAAGTGGC 214
|||||
QY 2547 TGGAGCGCTTCGGCTATACCAACCACTTACCCCTTACCTCGCTCCCGCTGCTCG 2606
|||||
Db 213 TGGAGCGCTTCGGCTATACCAACCACTTACCCCTTACCTCGCTCCCGCTGCTCG 154
|||||
QY 2607 CCTACTGCACCTCCCGCGCTGTGCTCTCTCAACAGGTTCATCATGCGGTGCG 2662
|||||
Db 153 CCTACTGCACCTCCCGCGCTGTGCTCTCTCAACAGGTTCATCATGCGGTGCG 98
|||||
```

```
RESULT 10
CG0300337
LOCUS
DEFINITION
CG0300337 835 bp DNA linear GSS 25-AUG-2003
genomic survey sequence.
ACCESSION
CG0300337
VERSION
CG0300337.1 GI:34214551
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 835)
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
TITLE
Consortium for Maize Genomics
JOURNAL
Unpublished (2002)
COMMENT
Other_GSSs: CG2AK93TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..835
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0743P17"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 835;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2307 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTTGAACGGGTTCAAGATGCAC 2366
|
|
|
Db 373 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTTGAACGGGTTCAAGATGCAC 432
|
|
|
QY 2367 GCCCGGGTGGCGTCCGTGCTACTGCGATGCGAAGCGGGCGGCTTCAAGGGGTCCGCGC 2426
|
|
|
Db 433 GCCCGGGTGGCGTCCGTGCTACTGCGATGCGAAGCGGGCGGCTTCAAGGGGTCCGCGC 492
|
|
|
QY 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 2486
|
|
|
Db 493 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 552
|
|
|
QY 2487 TCTTCTTTCAGCCGSCACAGCCCCCTGCTGTACGGCTACAAGAGCGGCAACCTCAAGTGGC 2546
|
|
|
Db 553 TCTTCTTTCAGCCGSCACAGCCCCCTGCTGTACGGCTACAAGAGCGGCAACCTCAAGTGGC 612
|
|
|
QY 2547 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTCGCTCCCGTGTCTCG 2606
|
|
|
Db 613 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTCGCTCCCGTGTCTCG 672
|
|
|
QY 2607 CCTACTGCACCTCCCGCGCTGTGCTCTCCTACCGGCAAGTTTCATATGCCGTGCG 2662
|
|
|
Db 673 CCTACTGCACCTCCCGCGCTGTGCTCTCCTACCGGCAAGTTTCATATGCCGTGCG 728
|
|
|
RESULT 11
CG081564/c
LOCUS
DEFINITION
CG081564 867 bp DNA linear GSS 20-AUG-2003
genomic survey sequence.
ACCESSION
CG081564
VERSION
CG081564.1 GI:33963858
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 867)
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
TITLE
Maize Genomics Consortium
JOURNAL
Unpublished (2003)
COMMENT
Other_GSSs: PUJFC40TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..867
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0654G07"
/clone_lib="ZM 0.6 1.0 KB"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 867;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2307 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTTGAACGGGTTCAAGATGCAC 2366
|
|
|
Db 411 AGCTGGGGTGGATCTACGGGTGCGATCAGGAGGACATCTTGAACGGGTTCAAGATGCAC 352
|
|
|
QY 2367 GCCCGGGTGGCGTCCGTGCTACTGCGATGCGAAGCGGGCGGCTTCAAGGGGTCCGCGC 2426
|
|
|
Db 351 GCCCGGGTGGCGTCCGTGCTACTGCGATGCGAAGCGGGCGGCTTCAAGGGGTCCGCGC 292
|
|
|
QY 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 2486
|
|
|
Db 291 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGGTCCGTCGAGA 232
|
|
|
QY 2487 TCTTCTTTCAGCCGSCACAGCCCCCTGCTGTACGGCTACAAGAGCGGCAACCTCAAGTGGC 2546
|
|
|
Db 231 TCTTCTTTCAGCCGSCACAGCCCCCTGCTGTACGGCTACAAGAGCGGCAACCTCAAGTGGC 172
|
|
|
QY 2547 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTCGCTCCCGTGTCTCG 2606
|
|
|
Db 171 TGGAGCGCTTGGCTATACACACACCATCTACCCCTTACCTCGCTCCCGTGTCTCG 112
|
|
|
QY 2607 CCTACTGCACCTCCCGCGCTGTGCTCTCCTACCGGCAAGTTTCATATGCCGTGCG 2662
|
|
|
Db 111 CCTACTGCACCTCCCGCGCTGTGCTCTCCTACCGGCAAGTTTCATATGCCGTGCG 56
|
|
|
RESULT 12
CG263400/c
LOCUS
DEFINITION
CG263400 917 bp DNA linear GSS 25-AUG-2003
genomic survey sequence.
ACCESSION
CG263400
VERSION
CG263400.1 GI:34175541
KEYWORDS
GSS.
```

SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 917)
TITLE Whiteclaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
JOURNAL Consortium for Maize Genomics
COMMENT Unpublished (2002)
Other GSSs: OG2AK93TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES Location/Qualifiers
source 1..917
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 917;
Best Local Similarity 100.0%; Pred. No. 6.3e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2307 AGCTGGGGTGGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 2366
DB 571 AGCTGGGGTGGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 512
QY 2367 GCGCGGGTGGCTCCGGTGTACTGATCCGAAGCGGGGGTTCAGAGGGTGGCGC 2426
DB 511 GCGCGGGTGGCTCCGGTGTACTGATCCGAAGCGGGGGTTCAGAGGGTGGCGC 452
QY 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCGGCTGGGGTTCGGTGCAGA 2486
DB 451 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCGGCTGGGGTTCGGTGCAGA 392
QY 2487 TCTTCTTACCGGGACAGCCCGCTGCTGACGGCTACAAGAACGGCAACCTCAAGTGGC 2546
DB 391 TCTTCTTACCGGGACAGCCCGCTGCTGACGGCTACAAGAACGGCAACCTCAAGTGGC 332
QY 2547 TGGAGGGCTTGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTCG 2606
DB 331 TGGAGGGCTTGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTCG 272
QY 2607 CCTACTGCACCCCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 2662
DB 271 CCTACTGCACCCCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 216
RESULT 13
CG300346/c
LOCUS CG300346 945 bp DNA linear GSS 25-AUG-2003
DEFINITION OG2AK93TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0743P17,
genomic survey sequence.
ACCESSION CG300346
VERSION CG300346.1 GI:34214560
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 818)
TITLE Whiteclaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
JOURNAL Consortium for Maize Genomics
COMMENT Unpublished (2002)
Other GSSs: OG2AK93TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES Location/Qualifiers
source 1..917
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 917;
Best Local Similarity 100.0%; Pred. No. 6.3e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2307 AGCTGGGGTGGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 2366
DB 820 AGCTGGGGTGGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 761
QY 2367 GCGCGGGTGGCTCCGGTGTACTGATCCGAAGCGGGGGTTCAGAGGGTGGCGC 2426
DB 760 GCGCGGGTGGCTCCGGTGTACTGATCCGAAGCGGGGGTTCAGAGGGTGGCGC 701
QY 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCGGCTGGGGTTCGGTGCAGA 2486
DB 700 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCGGCTGGGGTTCGGTGCAGA 641
QY 2487 TCTTCTTACCGGGACAGCCCGCTGCTGACGGCTACAAGAACGGCAACCTCAAGTGGC 2546
DB 640 TCTTCTTACCGGGACAGCCCGCTGCTGACGGCTACAAGAACGGCAACCTCAAGTGGC 581
QY 2547 TGGAGGGCTTGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTCG 2606
DB 580 TGGAGGGCTTGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTCG 521
QY 2607 CCTACTGCACCCCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 2662
DB 520 CCTACTGCACCCCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 465
RESULT 14
CC446974/c
LOCUS CC446974 818 bp DNA linear GSS 20-MAY-2003
DEFINITION FUHLI43TB ZM 0.6_1.0 KB Zea mays genomic clone ZMMBta477H13,
genomic survey sequence.
ACCESSION CC446974
VERSION CC446974.1 GI:30954665
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 818)
TITLE Whiteclaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
JOURNAL Consortium for Maize Genomics
COMMENT Unpublished (2002)
Other GSSs: OG2AK93TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES Location/Qualifiers
source 1..945
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBma0743P17"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 945)
TITLE Whiteclaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
JOURNAL Consortium for Maize Genomics
COMMENT Unpublished (2002)
Other GSSs: OG2AK93TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES Location/Qualifiers
source 1..945
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBma0743P17"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN
Query Match 10.3%; Score 356; DB 9; Length 945;
Best Local Similarity 100.0%; Pred. No. 6.3e-156;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2307 AGCTGGGGTGGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 2366
DB 820 AGCTGGGGTGGATCTACGGGTGATCAGGAGGACATCTCTGACGGGGTTCAAGATGCAT 761
QY 2367 GCGCGGGTGGCTCCGGTGTACTGATCCGAAGCGGGGGTTCAGAGGGTGGCGC 2426
DB 760 GCGCGGGTGGCTCCGGTGTACTGATCCGAAGCGGGGGTTCAGAGGGTGGCGC 701
QY 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCGGCTGGGGTTCGGTGCAGA 2486
DB 700 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCGGCTGGGGTTCGGTGCAGA 641
QY 2487 TCTTCTTACCGGGACAGCCCGCTGCTGACGGCTACAAGAACGGCAACCTCAAGTGGC 2546
DB 640 TCTTCTTACCGGGACAGCCCGCTGCTGACGGCTACAAGAACGGCAACCTCAAGTGGC 581
QY 2547 TGGAGGGCTTGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTCG 2606
DB 580 TGGAGGGCTTGGCTACATCAACACCAACATCTACCCCTTACCTCGCTCCGCTGCTCG 521
QY 2607 CCTACTGCACCCCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 2662
DB 520 CCTACTGCACCCCTCCCGCGGTGCTGCTCTCTCACCGGCAAGTTTCATCATCGCGTCG 465
RESULT 14
CC446974/c
LOCUS CC446974 818 bp DNA linear GSS 20-MAY-2003
DEFINITION FUHLI43TB ZM 0.6_1.0 KB Zea mays genomic clone ZMMBta477H13,
genomic survey sequence.
ACCESSION CC446974
VERSION CC446974.1 GI:30954665
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 818)
TITLE Whiteclaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
JOURNAL Consortium for Maize Genomics
COMMENT Unpublished (2002)
Other GSSs: OG2AK93TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES Location/Qualifiers
source 1..945
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBma0743P17"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUHLI437D
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source Location/Qualifiers
1..818

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa477H13"
/clone_lib="ZM 0.6-1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Query Match 10.1%; Score 349; DB 8; Length 818;
Best Local Similarity 100.0%; Pred. No. 1.3e-152;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1081 CAGGTACGAGAGGAGGAGCGTCTGCTGCGCGGTGGACCTGTTCTGAGGAC 1140
Db 357 CAGGTACGAGAGGAGGAGCGTCTGCTGCTGCGCGGTGGACCTGTTCTGAGGAC 298
Qy 1141 GGTGGACCCGCTCAAGGAGCGCGCTGTGACCGCAACCGTCTCTCCATCCTCGC 1200
Db 297 GGTGGACCCGCTCAAGGAGCGCGCTGTGACCGCAACCGTCTCTCCATCCTCGC 238
Qy 1201 CGTAGACTACCCGTGACCAAGGTCTCTGCTAGCTTCCGACGCGCGGTGATGCT 1260
Db 237 CGTAGACTACCCGTGACCAAGGTCTCTGCTAGCTTCCGACGCGCGGTGATGCT 178
Qy 1261 GACGTTTCGAGTCGCTGTCGAGAGCGCGCGAGTTCGCGCAAGTGGTGCCTTCTGCAA 1320
Db 177 GACGTTTCGAGTCGCTGTCGAGAGCGCGCGAGTTCGCGCAAGTGGTGCCTTCTGCAA 118
Qy 1321 GAAGTTCGGCATCGAGCCCGCGCGCGAGTTCCTGCTCAAGGTCGACTACCT 1380
Db 117 GAAGTTCGGCATCGAGCCCGCGCGCGAGTTCCTGCTCAAGGTCGACTACCT 58
Qy 1381 CAAGGACAAAGTGCAGCCACCTTCGTGACGAGCGCGCGCATGAAG 1429
Db 57 CAAGGACAAAGTGCAGCCACCTTCGTGACGAGCGCGCGCATGAAG 9

RESULT 15
CC56950
LOCUS OGLAD37TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0307H02,
DEFINITION genomic survey sequence.
GSS 1 (bases 1 to 452)

ACCESSION CC56950
VERSION CC56950.1 GI:32060243
KEYWORDS GSS.

SOURCE Zea mays
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 452)
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.

CONSORTIUM for Maize Genomics
Unpublished (2002)
TITLE Consortium for Maize Genomics
COMMENT Other_GSSs: OGLAD37TH

Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source Location/Qualifiers
1..452

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0307H02"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.6%; Score 332; DB 9; Length 452;
Best Local Similarity 100.0%; Pred. No. 1.3e-144;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1081 CAGGTACGAGAGGAGGAGCGTCTGCTGCGCGGTGGACCTGTTCTGAGGAC 1140
Db 121 CAGGTACGAGAGGAGGAGCGTCTGCTGCTGCGCGGTGGACCTGTTCTGAGGAC 180
Qy 1141 GGTGGACCCGCTCAAGGAGCGCGCTGTGACCGCAACCGTCTCTCCATCCTCGC 1200
Db 181 GGTGGACCCGCTCAAGGAGCGCGCTGTGACCGCAACCGTCTCTCCATCCTCGC 240
Qy 1201 CGTAGACTACCCGTGACCAAGGTCTCTGCTAGCTTCCGACGCGCGGTGATGCT 1260
Db 241 CGTAGACTACCCGTGACCAAGGTCTCTGCTAGCTTCCGACGCGCGGTGATGCT 300
Qy 1261 GACGTTTCGAGTCGCTGTCGAGAGCGCGCGAGTTCGCGCAAGTGGTGCCTTCTGCAA 1320
Db 301 GACGTTTCGAGTCGCTGTCGAGAGCGCGCGAGTTCGCGCAAGTGGTGCCTTCTGCAA 360
Qy 1321 GAAGTTCGGCATCGAGCCCGCGCGCGAGTTCCTGCTCAAGGTCGACTACCT 1380
Db 361 GAAGTTCGGCATCGAGCCCGCGCGCGAGTTCCTGCTCAAGGTCGACTACCT 420
Qy 1381 CAAGGACAAAGTGCAGCCACCTTCGTGCGAGG 1412
Db 421 CAAGGACAAAGTGCAGCCACCTTCGTGCGAGG 452

RESULT 16
CA183623

LOCUS CA183623
DEFINITION SCQGST3154G05.9 ST3 Saccharum officinarum cDNA clone SCQGST3154G05
5', mRNA sequence.

ACCESSION CA183623
VERSION CA183623.1 GI:35120577
KEYWORDS EST.

SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum

REFERENCE 1 (bases 1 to 608)
AUTHORS Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P

CONTRATO de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089

Email: parruda@unicamp.br
Clone distribution: Clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>
Plate: 154 row: G column: 05
Seq primer: T7 Promoter Primer.
Location/Qualifiers

FEATURES
source

1. .608
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQST3154G05"
/lab_host="DH10B"
/clone_lib="ST3"
/notes="Organ: Fourth apical stalk internodes of adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Fourth apical stalk internodes of adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 9.3%; Score 320; DB 6; Length 608;
Best Local Similarity 100.0%; Pred. No. 5.8e-139; Indels 0; Gaps 0;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 TCGCACAAACCGGAACGAGCTGGTGTGATCCGGGGCCACGAGGACCCCAAGCGCTGCGG 181
Db 152 TCGCACAAACCGGAACGAGCTGGTGTGATCCGGGGCCACGAGGACCCCAAGCGCTGCGG 211

QY 182 GCCTGAGCGGGCAGGTGTCGAGATATCGGCGACGAGGTCCGGCTCAGGTGGAGCGC 241
Db 212 GCCTGAGCGGGCAGGTGTCGAGATATCGGCGACGAGGTCCGGCTCAGGTGGAGCGC 271

QY 242 GACCTCTTCCTGCTGCAACGAGTGGCTTCCCGTGTGCGGCGCTGCTACGAGTAC 301
Db 272 GACCTCTTCCTGCTGCAACGAGTGGCTTCCCGTGTGCGGCGCTGCTACGAGTAC 331

QY 302 GAGCGCGGGAGGCGACGCAAACTGCCCCATGCAAGACGCGCTCAAGCGCCTCAAG 361
Db 332 GAGCGCGGGAGGCGACGCAAACTGCCCCATGCAAGACGCGCTCAAGCGCCTCAAG 391

QY 362 GGGAGCCCGAGGCTTCCGGGGACGATGACGAGGAGGACATCGACGACCTGGAGCGAG 421
Db 392 GGGAGCCCGAGGCTTCCGGGGACGATGACGAGGAGGACATCGACGACCTGGAGCGAG 451

QY 422 TTCAACATCGACGACGAGAA 441
Db 452 TTCAACATCGACGACGAGAA 471

RESULT 17
CA173049 624 bp mRNA linear EST 24-SEP-2003
LOCUS SCUTSB1033C12.g SBI Saccharum officinarum cDNA clone SCUTSB1033C12
5', mRNA sequence.
CA173049
CA173049.1 GI:35099270
EST.

ORGANISM
Saccharum officinarum
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
1 (bases 1 to 624)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST

JOURNAL
COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>
Plate: 033 row: C column: 12
Seq primer: T7 Promoter Primer.
Location/Qualifiers

FEATURES
source

1. .624
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCUTSB1033C12"
/lab_host="DH10B"
/clone_lib="SBI"
/notes="Organ: Stalk Bark from adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Stalk Bark from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 9.3%; Score 320; DB 6; Length 624;
Best Local Similarity 100.0%; Pred. No. 5.8e-139; Indels 0; Gaps 0;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 TCGCACAAACCGGAACGAGCTGGTGTGATCCGGGGCCACGAGGACCCCAAGCGCTGCGG 181
Db 125 TCGCACAAACCGGAACGAGCTGGTGTGATCCGGGGCCACGAGGACCCCAAGCGCTGCGG 184

QY 182 GCCTGAGCGGGCAGGTGTCGAGATATCGGCGACGAGGTCCGGCTCAGGTGGAGCGC 241
Db 185 GCCTGAGCGGGCAGGTGTCGAGATATCGGCGACGAGGTCCGGCTCAGGTGGAGCGC 244

QY 242 GACCTCTTCCTGCTGCAACGAGTGGCTTCCCGTGTGCGGCGCTGCTACGAGTAC 301
Db 245 GACCTCTTCCTGCTGCAACGAGTGGCTTCCCGTGTGCGGCGCTGCTACGAGTAC 304

QY 302 GAGCGCGGGAGGCGACGCAAACTGCCCCATGCAAGACGCGCTCAAGCGCCTCAAG 361
Db 305 GAGCGCGGGAGGCGACGCAAACTGCCCCATGCAAGACGCGCTCAAGCGCCTCAAG 364

QY 362 GGGAGCCCGAGGCTTCCGGGGACGATGACGAGGAGGACATCGACGACCTGGAGCGAG 421
Db 365 GGGAGCCCGAGGCTTCCGGGGACGATGACGAGGAGGACATCGACGACCTGGAGCGAG 424

QY 422 TTCAACATCGACGACGAGAA 441
Db 425 TTCAACATCGACGACGAGAA 444

RESULT 18
CG436514
LOCUS

DEFINITION CG436514 808 bp DNA linear GSS 17-SEP-2003
CG5EK86TV.ZM.0.7_1.5_KB_Zea_mays_genomic_clone_ZM5Mwa0846F03,
genomic survey sequence.
CG436514
CG436514.1 GI:34813053
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1 (bases 1 to 808)

REFERENCE
AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

TITLE
JOURNAL
COMMENT

Unpublished (2002)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

FEATURES
source

1..808
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0846P03"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBC5K; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.2%; Score 318; DB 9; Length 808;
Best Local Similarity 100.0%; Pred. No. 5e-138;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GGGTCGAGGCGGAGCTAGCTAGCAGCGCGGCTCCGCGCGATGGAGGCGAGCGC 105

Db 1 GGGTCGAGGCGGAGCTAGCTAGCAGCGCGGCTCCGCGCGATGGAGGCGAGCGC 60

Qy 106 CGGCTGTGTGCGCGCTCGCACACCGGAGAGCTGTGTGATCGGGGCGACGAGGA 165

Db 61 CGGCTGTGTGCGCGCTCGCACACCGGAGAGCTGTGTGATCGGGGCGACGAGGA 120

Qy 166 CCCCAGCGCTGCGGGCGCTGAGCGGCGAGTGTGCGAGATATGCGGCGACGAGTCCG 225

Db 121 CCCCAGCGCTGCGGGCGCTGAGCGGCGAGTGTGCGAGATATGCGGCGACGAGTCCG 180

Qy 226 GCTCACGCTGACGCGGACCTCTTGTGCTGCTGCAACAGAGTGGCGCTTCCCGTGGCG 285

Db 181 GCTCACGCTGACGCGGACCTCTTGTGCTGCTGCAACAGAGTGGCGCTTCCCGTGGCG 240

Qy 286 GCGCTGTACAGTACGAGCGCGGAGGCGACGAGACTGCCCCCGTCAAGAGCGC 345

Db 241 GCGCTGTACAGTACGAGCGCGGAGGCGACGAGACTGCCCCCGTCAAGAGCGC 300

Qy 346 CTACAAGCGCTCAAGGG 363

Db 301 CTACAAGCGCTCAAGGG 318

RESULT 19

CG454501

LOCUS

DEFINITION

CG454501

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 748)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and

Unpublished (2002)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..808

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBMA0846P03"

/clone_lib="ZM 0.7 1.5 KB"

/note="Vector: pBC5K; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

TITLE
JOURNAL
COMMENT

Unpublished (2003)

Other GSSs: PUIKA96TDB

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

FEATURES
source

1..748
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0601023"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Query Match 8.9%; Score 305; DB 9; Length 748;
Best Local Similarity 99.7%; Pred. No. 6.6e-132;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2307 AGCTGGGGTGGATCTACGGGTGCGATCACGAGGAGCATCTTACGGGGTTCAAGATGCACT 2366

Db 149 AGCTGGGGTGGATCTACGGGTGCGATCACGAGGAGCATCTTACGGGGTTCAAGATGCACT 208

Qy 2367 GCGCGGGTGGCGTCCGCTGTACTGTCATGCCAAGCGGGCGCGTTCAAGGGGTGCGCGC 2426

Db 209 GCGCGGGTGGCGTCCGCTGTACTGTCATGCCAAGCGGGCGCGTTCAAGGGGTGCGCGC 268

Qy 2427 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGTCCGTCGAGA 2486

Db 269 CGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGCGCTGGGTCCGTCGAGA 328

Qy 2487 TCTTCTTTCAGCGGACAGCGCCCTGTGTACGGCTACAAGACGCAACCTCAAGTGGC 2546

Db 329 TCTTCTTTCAGCGGACAGCGCCCTGTGTGTACGGCTACAAGACGCAACCTCAAGTGGC 388

Qy 2547 TGGAGCGCTTGGCTTACATCAACACCACTTACCCCTTACCTCGCTCCCGTGTCTCG 2606

Db 389 TGGAGCGCTTGGCTTACATCAACACCACTTACCCCTTACCTCGCTCCCGTGTCTCG 448

Qy 2607 CCTACTGCACCTCCCGCGCTGTGCTCTCTACCGGCAAGTTTCATATGCCGTCG 2662

Db 449 CCTACTGCACCTCCCGCGCTGTGCTCTCTCTACCGGCAAGTTTCATATGCCGTCG 504

RESULT 20

CC705263

LOCUS

DEFINITION

CG705263

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 640)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Unpublished (2002)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..748

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBTA0601023"

/clone_lib="ZM 0.6 1.0 KB"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high

Cot selected genomic DNA library"


```
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match          7.8%; Score 269; DB 8; Length 841;
Best Local Similarity 100.0%; Pred. No. 5.8e-115;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 817 GAACGACGAGCGAGCGAGCGCGTGTTCGAGGAAGTGTTCGATCGCGTCGAGCAAGGTGAA 876
Db 366 GAACGACGAGCGAGCGAGCGCGTGTTCGAGGAAGTGTTCGATCGCGTCGAGCAAGGTGAA 307

Qy 877 CCCGTACCGGATGTGTGATCGTGGTGTGCTGCTGTGTGCTGCTCTTCTTCTCCGCTACCG 936
Db 306 CCCGTACCGGATGTGTGATCGTGGTGTGCTGCTGTGTGCTGCTCTTCTTCTCCGCTACCG 247

Qy 937 TATCCTGCACCCCGCTCCCGGAGCGCATCGGGCTGTGGCTCTCCATCATCTCGAGAT 996
Db 246 TATCCTGCACCCCGCTCCCGGAGCGCATCGGGCTGTGGCTCTCCATCATCTCGAGAT 187

Qy 997 CTGGTTGGCCATCTCTCGATGCTCGACCAAGTTCCTCCCAAGTGTTCCTCCATCGACCGGA 1056
Db 186 CTGGTTGGCCATCTCTCGATGCTCGACCAAGTTCCTCCCAAGTGTTCCTCCATCGACCGGA 127

Qy 1057 GACGTACTCGACCGCTCTCCCTCAGGT 1085
Db 126 GACGTACTCGACCGCTCTCCCTCAGGT 98

RESULT 23
CC389096/c
LOCUS CC389096 651 bp DNA linear GSS 19-MAY-2003
DEFINITION PUHEM24TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBtra433C23,
genomic survey sequence.
ACCESSION CC389096
VERSION CC389096.1 GI:30869186
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 651)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUHEM24TB
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: Tg
Class: sheared ends.
FEATURES
source 1. 651
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBtra433C23"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN
Query Match          7.7%; Score 266; DB 8; Length 651;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 266; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 817 GAACGACGAGCGAGCGAGCGCGTGTTCGAGGAAGTGTTCGATCGCGTCGAGCAAGGTGAA 876
Db 366 GAACGACGAGCGAGCGAGCGCGTGTTCGAGGAAGTGTTCGATCGCGTCGAGCAAGGTGAA 307

Qy 877 CCCGTACCGGATGTGTGATCGTGGTGTGCTGCTGTGTGCTGCTCTTCTTCTCCGCTACCG 936
Db 306 CCCGTACCGGATGTGTGATCGTGGTGTGCTGCTGTGTGCTGCTCTTCTTCTCCGCTACCG 247

Qy 937 TATCCTGCACCCCGCTCCCGGAGCGCATCGGGCTGTGGCTCTCCATCATCTCGAGAT 996
Db 246 TATCCTGCACCCCGCTCCCGGAGCGCATCGGGCTGTGGCTCTCCATCATCTCGAGAT 187

Qy 997 CTGGTTGGCCATCTCTCGATGCTCGACCAAGTTCCTCCCAAGTGTTCCTCCATCGACCGGA 1056
Db 186 CTGGTTGGCCATCTCTCGATGCTCGACCAAGTTCCTCCCAAGTGTTCCTCCATCGACCGGA 127

Qy 1057 GACGTACTCGACCGCTCTCCCTCAGGT 1085
Db 126 GACGTACTCGACCGCTCTCCCTCAGGT 98

RESULT 24
CC705271/c
LOCUS CC705271 815 bp DNA linear GSS 19-JUN-2003
DEFINITION OGWEX08TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0562B16,
genomic survey sequence.
ACCESSION CC705271
VERSION CC705271.1 GI:32110047
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 815)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGWEX08TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: Tg
Class: sheared ends.
FEATURES
source 1. 815
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBma0562B16"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match          7.7%; Score 266; DB 9; Length 815;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 266; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1428 AGAGAGATGATGAGGAGTTCAGGTCGCGATCACGCGCTGGTGGCCCAAGCCCATGAGG 1487
Db 436 AGAGAGATGATGAGGAGTTCAGGTCGCGATCACGCGCTGGTGGCCCAAGCCCATGAGG 377

Qy 1488 AGAGAGATGATGAGGAGTTCAGGTCGCGATCACGCGCTGGTGGCCCAAGCCCATGAGG 1487
Db 436 AGAGAGATGATGAGGAGTTCAGGTCGCGATCACGCGCTGGTGGCCCAAGCCCATGAGG 377

Qy 1488 TGCCGCGAGAGGGGTGGATCATGAAGACGCGACGCGCTGGCCCGGGAACAACACCCGCG 1547
Db 370 TGCCGCGAGAGGGGTGGATCATGAAGACGCGACGCGCTGGCCCGGGAACAACACCCGCG 311

Qy 1548 ACACCCCGGATGATCATCAGGTGTTCCTGGGCGCACAGCGGGCGGCACGACACCGAGGGCA 1607
Db 310 ACACCCCGGATGATCATCAGGTGTTCCTGGGCGCACAGCGGGCGGCACGACACCGAGGGCA 251

Qy 1608 ACAGAGTCGCCCCGCTCGTGTACGCTCTCCGTGAGAGCGCGCGGATTCAGCAGCACCACA 1667
Db 250 ACAGAGTCGCCCCGCTCGTGTACGCTCTCCGTGAGAGCGCGCGGATTCAGCAGCACCACA 191

Qy 1668 AGAAGCGCGGCGCCATGAACGCTCTG 1693
Db 190 AGAAGCGCGGCGCCATGAACGCTCTG 165
```

```
QY 1548 ACCACCCCGCATGATCCAGGTGTTCTTGGGCGACAGCGCGGCCACGACACCGAGGGCA 1607
Db 316 ACCACCCCGCATGATCCAGGTGTTCTTGGGCGCACAGCGCGGCCACGACACCGAGGGCA 257
QY 1608 ACCAGCTGCCCGCTCGTCTAGCTCTCCCGTGAGAGCGCCCGGGATTCACGACCCACA 1667
Db 256 ACCAGCTGCCCGCTCGTCTAGCTCTCCCGTGAGAGCGCCCGGGATTCACGACCCACA 197
QY 1668 AGAAGCGCGCGCATGAAGCGCTCTG 1693
Db 196 AGAAGCGCGCGCATGAAGCGCTCTG 171

RESULT 25
LOCUS CC656942/c
DEFINITION OGLAD377H ZM 0.7_1.5_KB Zea mays genomic clone ZM5BMA0307H02,
genomic survey sequence.
ACCESSION CC656942
VERSION CC656942.1 GI:32060234
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 818)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGLAD37TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
Location/Qualifiers
1..818
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0307H02"
/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 7.7%; Score 266; DB 9; Length 818;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 266; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1428 AGAGAGTATGAGGAGTTCAAGTCCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 1487
Db 542 AGAGAGTATGAGGAGTTCAAGTCCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 483
QY 1488 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCGCTGTGGCCCGGGAAACAACACCCGCG 1547
Db 482 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCGCTGTGGCCCGGGAAACAACACCCGCG 423
QY 1548 ACCACCCCGCATGATCCAGGTGTTCTTGGGCGACAGCGCGGCCACGACACCGAGGGCA 1607
Db 422 ACCACCCCGCATGATCCAGGTGTTCTTGGGCGCACAGCGCGGCCACGACACCGAGGGCA 363
QY 1608 ACCAGCTGCCCGCTCGTCTAGCTCTCCCGTGAGAGCGCCCGGGATTCACGACCCACA 1667
Db 362 ACCAGCTGCCCGCTCGTCTAGCTCTCCCGTGAGAGCGCCCGGGATTCACGACCCACA 303
```

```
QY 1668 AGAAGCGCGCGCATGAAGCGCTCTG 1693
Db 302 AGAAGCGCGCGCATGAAGCGCTCTG 277

RESULT 26
LOCUS CG343883
DEFINITION OQ0AF68TV ZM 0.7_1.5_KB Zea mays genomic clone ZM5BMA0672K16,
genomic survey sequence.
ACCESSION CG343883
VERSION CG343883.1 GI:34261149
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 868)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OQ0AF68TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
Location/Qualifiers
1..868
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0672K16"
/clone_lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 7.7%; Score 266; DB 9; Length 868;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1428 AGAGAGTATGAGGAGTTCAAGTCCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 1487
Db 116 AGAGAGTATGAGGAGTTCAAGTCCGGATCAACGCGCTGTGGCCAGGCCATGAAGG 175
QY 1488 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCGCTGTGGCCCGGGAAACAACACCCGCG 1547
Db 176 TGCCGGCAGAGGGGTGGATCATGAAGGACGGCGCTGTGGCCCGGGAAACAACACCCGCG 235
QY 1548 ACCACCCCGCATGATCCAGGTGTTCTTGGGCGACAGCGCGGCCACGACACCGAGGGCA 1607
Db 236 ACCACCCCGCATGATCCAGGTGTTCTTGGGCGCACAGCGCGGCCACGACACCGAGGGCA 295
QY 1608 ACCAGCTGCCCGCTCGTCTAGCTCTCCCGTGAGAGCGCCCGGGATTCACGACCCACA 1667
Db 296 ACCAGCTGCCCGCTCGTCTAGCTCTCCCGTGAGAGCGCCCGGGATTCACGACCCACA 355
QY 1668 AGAAGCGCGCGCATGAAGCGCTCTG 1693
Db 356 AGAAGCGCGCGCATGAAGCGCTCTG 381

RESULT 27
LOCUS CG145361/c
795 bp DNA linear GSS 21-AUG-2003
```

```

DEFINITION PUJEW43TD.ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0652H13,
             genomic survey sequence.
ACCESSION   CG145361
VERSION     CG145361.1  GI:34036144
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 795)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
             Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
             Bennezen,J.
TITLE      Maize Genomics Consortium
JOURNAL     Unpublished (2003)
COMMENT     Other GSSs: PUJEW43TB
             Contact: Cathy Whitelaw
             TIGR
             9712 Medical Center Drive, Rockville, MD 20850, USA
             Tel: 301-838-5843
             Fax: 301-838-0208
             Email: whitelaw@tigr.org
             Seq primer: Tg
             Class: Sheared ends.
FEATURES   Location/Qualifiers
             source          1..795
             /organism="Zea mays"
             /mol_type="genomic DNA"
             /strain="B73"
             /db_xref="taxon:4577"
             /clone="ZMMBTa0652H13"
             /clone_lib="ZM 0.6-1.0 KB"
             /note="vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
             CoT selected genomic DNA library"
ORIGIN
Query Match      7.6%; Score 261; DB 9; Length 795;
Best Local Similarity 100.0%; Pred. No. 3.4e-111;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3146 TCCATTCTGCGGCTCCATCTTCTCCTGCTCTGGGTGAGGATCGACCTTTCATGCTC 3205
DB 795 TCCATTCTGCGGCTCCATCTTCTCCTGCTCTGGGTGAGGATCGACCTTTCATGCTC 736

QY 3206 AGGACCAAGGCGCGAGCTGAGGAGTGTGGCATCAATTGCTGAGCTGTTTATTAAAGT 3265
DB 735 AGGACCAAGGCGCGAGCTGAGGAGTGTGGCATCAATTGCTGAGCTGTTTATTAAAGT 676

QY 3266 TCAAAATTCTGAGCTTGTGCATAGGGAGAAAAAACAATTAGAAATTTTGAAGGTTG 3325
DB 675 TCAAAATTCTGAGCTTGTGCATAGGGAGAAAAAACAATTAGAAATTTTGAAGGTTG 616

QY 3326 TTGTGCTGTAATGTTATGTTACCAATGTCGACGAGGAATTGAACAAGGACAAG 3385
DB 615 TTGTGCTGTAATGTTATGTTACCAATGTTGACGAGGAATTGTCGACGAGGAATTGAACAAGGACAAG 556

QY 3386 GTTTGATTGTTAAATGGCAA 3406
DB 555 GTTTGATTGTTAAATGGCAA 535

RESULT 28
LOCUS      BE050992
DEFINITION ze71d11.g51 Maize Glume cDNAs Library Zea mays cDNA clone ze71d11
             5', mRNA sequence.
ACCESSION   BE050992
VERSION     BE050992.1  GI:8368047
KEYWORDS    EST.
SOURCE      Zea mays
ORGANISM    Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 437)
AUTHORS    O'Shaughnessy,A.L., Habermann,K., de la Bastide,M., Huang,E.N.,
             Nascimben,L.U., Schutz,K., Matero,A., Swaby,I., See,L.-H.,
             Preston,R.R., Rodriguez,M.A., Shah,R.S., Shekher,M., Spiegel,L.A.,
             Vil,M.D., Dedhia,N.N. and McComb,W.R.
TITLE      Expressed sequence tags from Zea mays (maize)
JOURNAL     Unpublished (2000)
COMMENT     Contact: W. Richard McComb
             Lita Annenberg Hazen Genome Sequencing Center
             Cold Spring Harbor Laboratory
             PO Box 100, Cold Spring Harbor, NY 11724, USA
             Tel: 516 367 8884
             Fax: 516 367 8874
             Email: mccomb@cschl.org
             Plate: za71 row: d column: 11
             Seq primer: -40M13RevUniv
             High quality sequence stop: 437.
FEATURES   Location/Qualifiers
             source          1..437
             /organism="Zea mays"
             /mol_type="mRNA"
             /db_xref="taxon:4577"
             /clone="za71d11"
             /clone_lib="Maize Glume cDNAs Library"
             /note="vector: Lambda Zap II (Stratagene); Site 1: XhoI;
             Site 2: EcoRI; Resistance: Ampicillin; Autoexcision:
             pBluescript SK (+/-); Titer: 7 x 10e-9 pfu/mL (as of
             9/28/94); Library source: John Doebley, B73"
ORIGIN
Query Match      7.1%; Score 246; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.1e-104;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1778 CGGAGGCCATGTGCTTCTCATGGACCTCAGTCGCGCGGAGGTCTGCTAGCTTCAG 1837
DB 192 CGGAGGCCATGTGCTTCTCATGGACCTCAGTCGCGCGGAGGTCTGCTAGCTTCAG 251

QY 1838 TTCCCGCAGAGTTCCGACGCGCATCGAGTCGACGACGATACGTAACAGGACACCGTC 1897
DB 252 TTCCCGCAGAGTTCCGACGCGCATCGAGTCGACGACGATACGTAACAGGACACCGTC 311

QY 1898 TTCTTCGACATCAACATGAAGGGCTGGACGGCATCCAAGCCCGGTGTAGTCGGGACA 1957
DB 312 TTCTTCGACATCAACATGAAGGGCTGGACGGCATCCAAGCCCGGTGTAGTCGGGACA 371

QY 1958 GGGTGGCTGTTCGGCGCCAGGCGCTCTACGGCTAGCAACCTCCCAAGGACCCAGAGG 2017
DB 372 GGGTGGCTGTTCGGCGCCAGGCGCTCTACGGCTAGCAACCTCCCAAGGACCCAGAGG 431

QY 2018 CCCAAG 2023
DB 432 CCCAAG 437

RESULT 29
LOCUS      BZ733703
DEFINITION OGEDM54TC.ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBMa0251J11,
             genomic survey sequence.
ACCESSION   BZ733703
VERSION     BZ733703.1  GI:28709833
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 821)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
             Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 12, 2005, 10:16:14 ; Search time 5122 Seconds
(without alignments)
7817.964 Million cell updates/sec

Title: US-10-627-132-30

Perfect score: 5648

Sequence: 1 MEASAGLVAGSHNRNELVLI.....IDPFIVRTKGPDPVRCQGINC 1052

Scoring table:

BLOSUM62
-Q=/cgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgapop 10.0 , Xgapext 0.5
-DB=EST -QFMT=fastcap -SUFFIX=est -MINMATCH=0.1 -FOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10627132 @CGN 1 1 3043 @runat 04032005 113218 21160 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5347.5	94.7	3168	9	CL976367
2	4089	72.4	3783	3	AY104730
3	4061.5	71.9	3897	3	AY103655
4	3938	69.7	3788	3	AY103701
5	3912.5	69.3	3763	3	AY108113
6	3907	69.2	3192	9	CL972423
7	3821.5	67.7	3898	3	AY110415
8	3719	65.8	3911	3	CNS0A1ML
9	3633.5	64.3	3728	3	AY112236

10	3597.5	63.7	3696	3	AY110079
11	3350.5	59.3	2872	3	AY104236
12	2525	44.7	1874	6	CD726831
13	2311	40.9	3474	9	CL978864
14	2134.5	37.8	2595	9	CL964957
15	1982	35.1	3048	9	CL974686
16	1607.5	28.5	2640	9	CL973331
17	1599.5	28.3	2670	9	CL973334
18	1549.5	27.4	2607	9	CL973335
19	1539.5	27.3	960	1	AB049602
20	1531	27.1	1797	9	CL979846
21	1434	25.4	898	7	CO089073
22	1432.5	25.4	1189	3	AY107656
23	1424	25.2	916	7	CK272603
24	1421	25.2	924	7	CF513822
25	1395	24.7	789	7	CF430024
26	1365	24.2	1152	3	CNS0A924
27	1362.5	24.1	925	7	CF213555
28	1344	23.8	953	6	CD439980
29	1305	23.1	848	7	CN129331
30	1302.5	23.1	1022	8	BZ827769
31	1275	22.6	821	9	CG383072
32	1272.5	22.5	853	7	CO086716
33	1272	22.5	875	7	CF514581
34	1268	22.5	869	7	CO089781
35	1264	22.4	779	6	CA916873
36	1262.5	22.4	837	7	CF512600
37	1262	22.3	834	7	CO519653
38	1258	22.3	792	6	CB671068
39	1249	22.1	772	6	CB646549
40	1246	22.1	790	6	CA161838
41	1246	22.1	851	6	CB892280
42	1240	22.0	799	6	CB894714
43	1239	21.9	819	7	CO111111
44	1237	21.9	710	6	CA261021
45	1236	21.9	868	9	CC6660677

ALIGNMENTS

RESULT 1

CL976367

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CL976367 3168 bp DNA linear GSS 21-SEP-2004
OsIFCC028793 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.

CL976367 GI:52407247

GSS.

Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 3168)

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Jiao, X., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,

Wong, G.K.S., Deng, X.W. and Wang, J.

An analysis of transcriptional regulation of the rice genome and

its comparison to Arabidopsis

Unpublished (2004)

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1..3168

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

```
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3168
Score: 5347.50 Matches: 992
Percent Similarity: 96.78% Conservative: 30
Best Local Similarity: 93.94% Mismatches: 29
Query Match: 94.68% Indels: 5
DB: 9 Gaps: 4

US-10-627-132-30 (1-1052) x CL976367 (1-3168)
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuLe 20
Db 1 ATGGAGGCGAGCGCGGCTGGTGGCGGGTTCGACCAACCGGAACGAGCTGGTCTGATC 60
Qy 21 ArgGlyHisGluAAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluLeu 40
Db 61 CGGGGGCAGCAGGAGCCCAACCGCTGGCGGCGCTGAGCGGCGAGGTGTGGAGATATGC 120
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 121 GCGCAGCAGGTCGCGCGCACCGTCGACGCGGCGACCTTCCTGCTCGCTGCAACGAGTGGCGGC 180
Qy 61 PheProValCysArgProCysGlyGluThrGluArgArgGluGlyThrGluAsnCysPro 80
Db 181 TTCGCGGTGTCCGCGCTGTACGAGTACGAGGCGCGCGGAGCCACCCAGAACTGGCCCC 240
Qy 81 GlnCysLeuThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 241 CAGTGCAAGACCCGCTACAGCGCTCAAGGGGAGCCCGAGGGTCCCGGGGACGAGAC 300
Qy 101 GluGluAepIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnGln 119
Db 301 GAGGAGGACATTGACGCTGGAGCACGAGTTCAACATCGACGACGAGAACGACGAGC 360
Qy 120 ---ArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 361 CTGCGACGAGGATCAGGATGGCATGCAGAACAGCCACATCACCGAGGCGATCTGCACGGC 420
Qy 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
Db 421 AAGATGACTACGGAGGGGGCCCCGACCGCGCGCGGCGGCAACAGCACCCCG---CTCCCG 477
Qy 159 ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
Db 478 CCGATCATCACCGCGCTCGCTCGCTCGGTCGAGCGGGGAGTTCCCGATATCGAACAGC 537
Qy 179 TyrGlyHisGlyGluValSerSerSerLeuHisLysArgIleHisProTyrProValSer 198
Db 538 CATGCCATTCGCGAGTTCTCTCTTCTCCCTGCACAAAGCGCATCCACCCCTACCCGCTGTCT 597
Qy 199 GluProGlySerAlaLysThrAspGluLysGlyValSerTrpLysGluArgMetAsp 218
Db 598 GAGCAGGAGGTGCAAGAGTGGACGAGAAAGAGGTGAGCTGGAGAGAGGATGGAC 657
Qy 219 AspTrpLysSerLysGlnGlyLeuGlyGlyAla---AspProGluAspMetAsp 237
Db 658 GACTGGAAATCCACGAGGGCATCGTCGCGCGCGCGCCCCCGATCCCGACGACTACGAC 717
Qy 238 AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla 257
Db 718 CGCGACGTCCTCACTGAACACGAGCGGAGCGCGCTGTCGAGGAAGGTGTGCGATCGCG 777
Qy 258 SerSerLysValAsnProTyrArgMetValIleValValArgLeuValValAlaPhe 277
Db 778 TCGACCAAGGTGAACCCGTACCGGATGGTGATCATCTCCCTCGCTCGTGGTCTCGGCTTC 837
Qy 278 PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSer 297
Db |||||
```

```
Db 838 TTCTCTCGGTACCGCATCTCTCCACCGGTGTCGCGACGCGCATCTCCGCTGTGGCTCACCTCC 897
Qy 298 IleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhe 317
Db 898 ATCATCTGCAGATCTGGTTCGCGCTGTGTGGATCTCTCACCAGTTCCTCCAAAGTGGTAC 957
Qy 318 ProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGlu 337
Db 958 CCATCGACAGGAGGAGACTTACCTCGACCGCTCTCTCTCTCGCTACGAGCGCGAGGGGAG 1017
Qy 338 ProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluPro 357
Db 1018 CCGTCTGCTATCTCGCGGTGGACCTGTCTCGTCAGCACGCTGGATTCGCTCAAGGAGCCG 1077
Qy 358 ProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys 377
Db 1078 CCGTGTGTACGCGCAACACCGTGTGTCTCTCTCGCTGCTACCTACCCGCTCGACAG 1137
Qy 378 ValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGlu 397
Db 1138 GTGTCTGTCTACGCTCTCCGACGCGCGCTCCATCTCAGTTCGAGTCTGTCTGGAG 1197
Qy 398 ThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArg 417
Db 1198 ACGCGGAGTTCGCGCGCAAGTGGTGCCATTTCTGCAAGAAGTTTCAGCATCGAGCCCGC 1257
Qy 418 AlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThr 437
Db 1258 GCGCGGAGTTCATTTCTCCGAGAGTGCATCTCAAGGACAGAGTTCATCCCAAC 1317
Qy 438 PheValGlnGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsn 457
Db 1318 TTCTGTCCAGGCGCGCGCCATGAAGAGAGTACGAGGAGTTCAAGGTGAGGATCAAC 1377
Qy 458 AlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThr 477
Db 1378 GCGTGTGGCGAAGCGCGAAGGTGCGCGGAGAGGTGGATCATGAAGAGCGGAGC 1437
Qy 478 ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 497
Db 1438 CCATGGCGGGGAAACAACCCCGCACCCCGGGCATGATCCAGGTGTCTCTCGGCCAC 1497
Qy 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 517
Db 1498 AGCGCGCGCCACGACACCGAGGGCAACGAGCTCCCGCTCCCTCGTCTACGTCTCCCTGAG 1557
Qy 518 LysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgVal 537
Db 1558 AAGCGCCCGGCTTCCAGCACCAAGAGCGCGCGCCATGAACGCGCTCATTTCTGTGTG 1617
Qy 538 SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn 557
Db 1618 TCGCGCTGTGTGACGAACGCGCGCTTCATGCTCAACTTGGATTGGATCACTATCATCAAC 1677
Qy 558 AsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
Db 1678 AACAGCAAGGCCATCAGGAGGCGATGTGTCTCTCATGGACCCCGAGGTCCGACGGAAG 1737
Qy 578 ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAla 597
Db 1738 GTGTGTCTACGTGCTTCCGAGGTTCCCGCAGAGTTCCAGCGCATCGACGCTCCAGCCGATACGCC 1797
Qy 598 AsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
Db 1798 AACCGCAACCCGCTTCTTCGACATCAACATGAAGGGGCTTGATGGGATCCAGGGGCCG 1857
Qy 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProPro 637
Db 1858 GTGTGTCTCGGACAGGCTGTGTTCAGGCGCGAGGCGCTGTATCGGATCAACACCTCTCC 1917
Qy 638 LysGlyProLysArgProLysMetValThrCysAspCysProCysPheGlyArgLys 657
Db 1918 AAGGACCCCAAGAGGCCCAAGATGTTGACCTGCGCATGCTGCTCCCTTGTCTCGGGAGGAG 1977
```

```
QY 658 IysArgLysHisAlaLysAspGlyLeuProGluGly---ThrAlaAspMetGlyValAsp 676
Db 1978 AAGCGGAAGCACGGCAAGACGGCTCCCGGAGGGCGTCCGCGCGACGGCGGATGGAC 2037
QY 677 SerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAla 696
Db 2038 AGCGACAGGAGATGCTCATGTCGAGATGAATCTCGAGAGCGGTTCGGGACGTGGCG 2097
QY 697 AlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAla 716
Db 2098 GCGTTCTGTCGACGTCGACGTCGATGAGGAGGCGGTCGCGCGCTCGTCCAGCCCGCC 2157
QY 717 AlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrp 736
Db 2158 GCGCTCTCTCAAGAGGCCATCATGTCATGCTGAGTCTGAGTACGAGCAAGACCGACTGG 2217
QY 737 GlyLeuGluLeuGlyTyrPheGlySerIleThrGluAspIleLeuThrGlyPheLys 756
Db 2218 GGTCTCGAGCTGGGTGGATCTACGGGTGATCAGGAGGATCTTAACTAACGGGGTTCAG 2277
QY 757 MetHisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGly 776
Db 2278 ATGCACCTGCGGGGTGGAGATCGGTGACTCATGTCATGCCGAGAGGGCGGTTCAGGGG 2337
QY 777 SerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySer 796
Db 2338 TCGGCGCGGATCAACCTATCTGACCGTCTCAACACGAGTCTCGGTGGGTCTCGGCTCC 2397
QY 797 ValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeu 816
Db 2398 GTGAGATCTCTTCAGCGCGGACAGCGCGCTCTTACGGGTATAGAACGCGCAACCTT 2457
QY 817 LysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuPro 836
Db 2458 AAGTGGCTCGAGCGTCTCTTACATCAACACCATCACTCCCTTCACTCTCCCTCC 2517
QY 837 LeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro 856
Db 2518 CTCTCTCGGCTACTGCACTCTCCCGCGCTCTGCTCTCCACCGCAAGTTCATCATGCT 2577
QY 857 SerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThr 876
Db 2578 CGATTACAGCTTTGCGAGTTGTCTTCATCGCGCTCTTCATCTCTCTCTCGGACG 2637
QY 877 GlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGln 896
Db 2638 GGCATCTCTGGAGATGAGTGGAGCGGGGTGAGCATCGAGGAGTGTGGAGGACGAGCAG 2697
QY 897 PheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLys 916
Db 2698 TTCTGGGTCTATCGCGGGGTGTTCGGCGCACCTGTTCGGGTGGTTCGAGGGCTCTCAAG 2757
QY 917 ValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAsp 936
Db 2758 GTGCTGGCGGGATCGACACCACTTCACCGTCAGTCCAGGCGCCACCGAGACGAGGAC 2817
QY 937 AspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThr 956
Db 2818 GACGAGTTTCGGGAGCTATACGCTTCAAGTGGACCACTCTCTCATCTCCACCCACGAC 2877
QY 957 LeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGly 976
Db 2878 CTGCTCATCTCAACATCATCGCGCTCTCGCGCGGTCTCGAGCGCCATCAACACGCG 2937
QY 977 TyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHis 996
Db 2938 TCCGAGGCGTGGGGCGGCTCTTCGGAAGCTCTTCTTCGCGCTCTCGGTTCATCGTCCAC 2997
QY 997 LeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgTrpProThrValValVal 1016
Db 2998 CTCTACCTCTCTCAAGGGGCTCATGGGAGGAGAACCGGACGCCCCACCATTTGTTGTC 3057
```

```
QY 1017 lleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProphe 1036
Db 3058 ATCTGGTCGGTCTGCTCGCTCCATCTTCTCTTCTGCTCGGTCAGGATTCATCCCTTC 3117
QY 1037 lleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3118 ACCATCAAGCCAGGGCGCTGACGTGAGCGAGTGGCATCACTGC 3165

RESULT 2
LOCUS AY104730 3783 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PCO100501 mRNA sequence.
ACCESSION AY104730
VERSION AY104730.1 GI:21207808
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3783)
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3783)
AUTHORS Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1. 3783
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 0 Length: 3783
Score: 4089.00 Matches: 760
Percent Similarity: 79.55% Conservative: 127
Best Local Similarity: 68.16% Mismatches: 138
Query Match: 72.40% Indels: 90
DB: 3 Gaps: 16
```

US-10-627-132-30 (1-1052) x AY104730 (1-3783)

```
QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 186 ATGAGGCGCAGCGCGGCTGTCGCGCTCGCACCAACCGGACGAGCTCGTCTCATC 245
QY 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 246 CGCCGCGATGCGAGCGCGGCGCGGAGCCCATGGCCAGCGGACCGGCGAGTGTGCCAG 305
QY 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
```


Qy	718	LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGly	737
Db	2454	CTTCTGAAGGAGCTATACATGTCTCATCAGCTGTGGATATGAAGACAACACAGCGTCGGGA	2513
Qy	738	LeuGluLeuGlyTpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMet	757
Db	2514	AAGAATTTGGTGTAATTATGGATTCAGTCAAGAAGATAATCTTACTGGGTAAAGATG	2573
Qy	758	HisCysArgGlyTpArgSerValTyrCysMetProLysArgAlaLaPheLysGlySer	777
Db	2574	CAC TGCCATG GTGGCGTCAATTTACTGCATACCTAAA CGGCGCGCTTCAAAGGTTC	2633
Qy	778	AlaProIleAenLeuSerAspArgLeuAenGlnValLeuArgTrpAlaLeuGlySerVal	797
Db	2634	GCACCTCTCAATCTTCCGATCGTCTTACCAGGTTCTTCGGTGGGCTCTTGGTTCATTT	2693
Qy	798	GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAenGlyAenLeuLys	817
Db	2694	GAATTTCTTCAGAACCACTGCCCTCTCTGGTATGGGTAT---GGTGGTGACATAAG	2750
Qy	818	TrpLeuGluArgPheAlaTyrIleAasnThrThrIleTyrProPheThrSerLeuProLeu	837
Db	2751	TTCCTGGAAAGGTTTCGTACATTAACCTCCATCGTATACCTCTGGACATCTATCCCGCTC	2810
Qy	838	LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer	857
Db	2811	TTGGCTATTGCACATTGCCTGCCATCTGCTGTGCTGACAGGAAATTTATCACGCCACAG	2870
Qy	858	IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly	877
Db	2871	CTTTAACCAATGTTGCCAGCGCTGTGGTTCATGTCACATTTTTTCATCTGTCATTTTTT	2930
Qy	878	IleLeuGluMetArgTpSerGlyValSerIleGluGluTrpTrpArgAenGluGlnPhe	897
Db	2931	ATCTCTGGAAATGAGATGGAGTGGTGTAGGCATCGATCAGCTGGTGGAGAAACGACGAGT	2990
Qy	898	TrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysVal	917
Db	2991	TGGGTCAATTGGAGCGGTCTTTCACATCTCTTTGCTGTGTCCAGGACCTCTCCAAGTCC	3050
Qy	918	LeuAlaGlyIleAspThrAenPheThrValThrSerLysAlaThrGlyAspGluAspAsp	937
Db	3051	ATAGCTGGTGPAGACACGAGCTTCACATGTGACATCCAAG-----GGCGGAGACGACG	3104
Qy	938	GluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeu	957
Db	3105	GAGTTCACAGAGCTGACACANTCAATATGACGACCCCTCTGTATPACCTCCGACAAACCTG	3164
Qy	958	LeuIleAenValIleGlyValValAlaGlyIleSerAspAlaIleAenAsnGlyTyr	977
Db	3165	CTCCTACTGAACCTCATTTGGAGTGGTAGCTGGCATCTCCAATGGATCAACACGGNAT	3224
Qy	978	GlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisLeu	997
Db	3225	GAATCATGGGGCCCCCTGTTCCGGAAAGCTCTTCTTTCATTTTGGGTGATCGTCCATCTT	3284
Qy	998	TyrProPheLeuLysGlyLeuMetGlyArgGlnAenArgHrpProThrValValIle	1017
Db	3285	TACCCGTTCTCAAGGGTCTGTTGGGAGGCAGAACAGGACGCCAACGATGTGCATGTGTC	3344
Qy	1018	TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgileAaspPropheIle	1037
Db	3345	TGGTCCATCTCTCCGGTTCGATCTTCTCGCTGCTTTGGTCCGGATCGACCCGCTTCCTT	3404
Qy	1038	ValArgThrLysGlyProAspValArgGlnCysGlyIleAenCys	1052
Db	3405	GCGAAGGATGATGCTCCCTGTTGGAGGAGTGTGCTGTGATTGC	3449
RESULT 3			
AY103655			
LOCUS	AY103655	3897 bp	mRNA linear HTC 16-OCT-2002
DEFINITION	Zea mays PCO096398 mRNA sequence.		

[illegible]

1467	TTGAAAGACAGGTGGCGCAAACTTTGTTAGGGACAGAGAGCAATGAAGAGAGNGTAT	1526
450	GlulGluphelysValArgIleAenAlaLeuValAlaLysAlaMetIysValProAlaGlu	469
1527	GAGGAATTCAGGTGAGATCAATGCTTAGTTGGCCAAAGCCAGAAAGTTCTCTGAAGAA	1586
470	GlyTrpIleMetIysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly	489
1587	GGATGTGACAAATGAAGATGAACCCCTCGCTGGAAACAATGTTCTGTATCATCTCTGA	1646
490	MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAenGluLeuPro	509
1647	ATGATTCAGGTCCTCTCTGGCCAAAGCGGAGGCTTGACTGTGAGGGAATGAATGAC	1706
510	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisIshIysLysAlaGly	529
1707	CGATTGGTTATGTTCTAGAGAGAAACGACAGGCTATAACCATCATAAAGAAAGCTGT	1766
530	AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn	549
1767	GCTATGAATGCATTTGGTCCGAGTCTCTGCTGTACTAACCAATGCTCCATATTGTTAAAC	1826
550	LeuAspCysAspHisIstYrIleAenAsnSerLysAlaIleArgGluAlaMetCysPheLeu	569
1827	TTGGATTGTGATCACATACATCAACACAGCAAGGCTATAAGAGAGCAATGTGTTATG	1886
570	MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly	589
1887	ATGGACCTTTACTAGGAAGAGAGGTTTGCTATGTACAGTTCCCTCAAGAATTTGATGG	1946
590	IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys	609
1947	ATTGATCGCCATGACCGATATGCTAAACCGGAATGTTGTCTTTTTCATATCAACATGAA	2006
610	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln	629
2007	GGTTTGGATGGTATTTCAGGTGCAAAATATGTTGGTACGTGTGTATTTAGAGGCAG	2066
630	AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp	649
2067	GCATTATATGGTTATGATGCCCCCANA--ACAAGAAGCCACCATCAAGGACTTGCAAC	2123
650	-----CysCysProCysPheGlyArgLysIysArgLysHisAla	662
2124	TGCTGGCCCAAGTGGTCTTTTGTCTGTCTGTTGGCAATAGGAAGCAAAAGAGACT	2183
663	LysAspGlyLeuProGlu-----	668
2184	ACCAAAACCCAAACAGAGAGAAAGATTAATTTTTCAGAAAGAGAGAACCAATCC	2243
669	-----GlyThrAlaAspMetGlyValAspSerAspLysGlu	680
2244	CCTGCATATGCTCTTGTGTGAAATGACGAAGCTCTCCAGGAGCTGAGATGAAGGCC	2303
681	MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr	700
2304	GGTATTGTAATCAACAAAAATTAGAAAGAAATTTGGCCAACTCTCTGTTTGTGTACA	2363
701	SerThrLeuMetGluGluGlyGlyValProProSerSerSerProAlaAlaLeuLeuLys	720
2364	TCCACACTTCGAGAAATGTGGAAACCTTGAAGAGTGCAGTCTCTCTCTTTTGA	2423
721	GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu	740
2424	GAAGCTATACATGTCATTAGTTTGTGGTTATGAAGACAAGACAGACTGGGGAAAGAGATT	2483
741	GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg	760
2484	GGCTGGATCTATGATCAGTTCTACAGAAGATATTCTAATCTGTTTCAAGATGCAATTGT	2543
761	GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle	780
2544	GGTTGGCGGTCAATTTACTGATCACTAAACGGGTTGCATTTCAAGGTTCTGCACCTCT	2603

QY 781 AsnLeuSerAspArgLeuGlnValLeuArgTrpAlaLeuGlySerValGluLeuPhe 800
 Db 2604 AATCTTTCAGATCGCTTCCACGAGTCTTGGGGCTTGGGCTATAGATCTTC 2663
 QY 801 PheSerArgHisSerProLeuLeuTrpGlyLysAsnGlyAsnLeuLysTrpLeuGlu 820
 Db 2664 TTCAGCAATCATGCCCTCTTGGTAAGGTAT---GGTGGCGGTCTGAATTTTGGAA 2720
 QY 821 ArgPheAlaTrpIleAsnThrThrIleTrpProPheThrSerLeuProLeuLeuAlaTrp 840
 Db 2721 AGATTTTCTACATCACTCCATCGTGTATCTTGGACATCTATCCCTCTTGGCTTAC 2780
 QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
 Db 2781 TGTACATTCCTCCATCTGTTTATTGACAGGAAATTTATCACTCCAGAGCTGAATAAT 2840
 QY 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyLeuGlu 880
 Db 2841 GTTCCAGGCTGTGGTTCATGTCACCTTTTATCTGCAATTTTGTACGAGCATCTTAGAA 2900
 QY 881 MetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIle 900
 Db 2901 ATGAGATGGAGTGGTTCGAATTCATGACTGGTGGAGGAATGAGCAGTCTGGGTCAAT 2960
 QY 901 GlyGlyValSerAlaHisLeuPheAlaValAlaValGlnGlyLeuLysValLeuAlaGly 920
 Db 2961 CGAGGTGTCTCTCACCTCTTGTCTGTGTCCAGGGACTTCTCAAGGTCTAGCTGT 3020
 QY 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
 Db 3021 GATGATACAGCTTCCAGCTGACATCAAG-----GGTGGAGATGATGAGGATCTCA 3074
 QY 941 GluLeuTrpAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleIle 960
 Db 3075 GAGCTATATACATTCAAATGAGTACCTTATTGATACCTCTTACCACCTTGTCTTATG 3134
 QY 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTrpGlnSerTrp 980
 Db 3135 AACTTCATTTGGTGTCTCTGGCTTCAATGCGATCAATCAATGATGATGATGATGATG 3194
 QY 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTrpPhe 1000
 Db 3195 GCGCCCTCTTGGGAAGCTATCTTGTCAATTTGGGTGATTTGATCTTATCCCTTT 3254
 QY 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgTrpProThrValValIleTrpSerIle 1020
 Db 3255 CTCAAAGGTTTGGTGGAGGCAAAACAGACACCAAGATTTGATCTGTCTGTCTCAAT 3314
 QY 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
 Db 3315 CTGCTGGCTTCAATCTTCTCGCTCTTGGGTTCGATTTGATCTTCTCTCGAAGAT 3374
 QY 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db 3375 GATGTCCTGCTTCTTGGAGGATGTTGGATTGCT 3410

RESULT 4
 AY103701
 LOCUS AY103701 3788 bp mRNA linear HTC 16-OCT-2002
 DEFINITION Zea mays PCO120363 mRNA sequence.
 ACCESSION AY103701
 VERSION AY103701.1 GI:21206779
 KEYWORDS HTC.

SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD;
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 3788)
 Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 Maize Mapping Project/DuPont Consensus Sequences for Design of

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 3788)
 Coe,E.H.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 If you are interested in getting corresponding physical clones,
 these are publicly available from ZmDB. www.zmdb.iastate.edu; TIGR,
 searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZmDB:
 www.zmdb.iastate.edu.

FEATURES
 source

1..3788
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="MaizeDB:637005"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3788
 Score: 3938.00 Matches: 751
 Percent Similarity: 78.26% Conservative: 120
 Best Local Similarity: 67.48% Mismatches: 161
 Query Match: 69.72% Indels: 81
 Gaps: 17

US-10-627-132-30 (1-1052) x AY103701 (1-3788)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
 Db 201 ATGAGGCTAGCGGGGCTGGTCCGGCTCGCATACCGGAACGAGCTGTGTGTATC 260
 QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnVal----- 36
 Db 261 CGC-----CGCAGCCGAGTCCGGAGCCCGGGCGCGCGCGCGCGCGGAG 314
 QY 37 -----CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheVal 54
 Db 315 GCGCGTCCAGATATCGCGACGAGTCCGGGTGGGCTTCGACCGGGAGCCCTTCGTG 374
 QY 55 AlaCysAsnGluCysGlyPheProValCysArgProCysTrpGlyTrpGluArgGlu 74
 Db 375 GCGTGCACAGTGCCTTCCCGTCTGCGCGCGCTGTCTACGAGTACGAGCGCGCGAG 434
 QY 75 GlyThrGlnAsnCysProGlnCysLysThrArgTrpLysArgLeuLysGlySerProArg 94
 Db 435 GGCTCGAGGCGTCCCGCAGTGCAGGACCCGCTACAGCGCTCAAGGGTTCGCCGGG 494
 QY 95 ValAlaGlyAspAspGluGluAspIleAspLeuGluHisGluPheAsnIleAsp 114
 Db 495 GTGGCTGGGACGAGGAGGAGCGGCTCGACGACCTGGAGGGGAGTTCGGCCTGCGAG 554
 QY 115 AspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGln---IleThrGlu 133
 Db 555 GACGGC-----GCCGGCCACGAGGACGCCCGCAGTACGTCCCGCAG 596
 QY 134 AlaMetLeuHisGlyArgMetSerTrpGlyArgGlyProAsp---AspGlyAspGlyAsn 152
 Db 597 TCCATGTCTGGGCGCAGATGATGATCGCCCGCGCGCGCGCGCGCGCGAGCTTCAAC 656
 QY 153 AsnThrProGlnIleProProIleIleThrGlySerArgSerValProValSerGlyGlu 172

Db	657	CCCGTCCCAACGTCGGCGCTCCTCACCAACGGCCAGATGGTTGATCAGATCCCGCGGAG	716
Qy	173	-----PheProIleThrAenGlyTyGlyHisGlyGluValSerSerLeu	188
Db	717	CAGCAGCGCTCGTGGCGTCTCATGCGGGCGGGCGCGCGC	758
Qy	189	HisLysArgIleHisProTyrProValSerGluProGly-----SerAlaLys	204
Db	759	GGCAGAGGATCACCGCTCCCTTCGTGTATCCAACTTCCAGTGCACCGAGATCC	818
Qy	205	TrpAspGluLysLysGlu-----ValSerTrpLysGluArg	216
Db	819	ATGACCGCTCCAAGGATCTGGCGCCTACGGATATGGCAGCGTGGCTCGAAGGAGAGA	878
Qy	217	MetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGlyAla	231
Db	879	ATGGAGGCTGGAAAGACAGAGAGCGCTCGAGCATGTCAGGACGAGGGTGGCGGT	938
Qy	232	AspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSer	251
Db	939	GATTGGGATGGCAGCATGCAGATCTGCCACTAATGGATGAAGTAGGCAGCCATTGTCC	998
Qy	252	ArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleValValArg	271
Db	999	AGAAAGTCCCTATATCATCAAGCCGAATAATCCCTACAGGATGATATCGTTATCCCG	1058
Qy	272	LeuValValIleAlaPhePheLeuArgTyArgIleLeuHisProValProAspAlaIle	291
Db	1059	TTGGTGGTTTGGGTTCCTCTCCACTACCGAGTGATGATCCGGGAAGATGCATTT	1118
Qy	292	GlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAsp	311
Db	1119	GCATTGGCTCATATCTGAATCTGGAATCTGGTTCCGATGTCCTGGATCTTGAT	1178
Qy	312	GlnPheProLysTrpPheProIleAspArgGluThrTyTrpLeuAspArgLeuSerLeuArg	331
Db	1179	CAGTCCCAAGTGGCTTCCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGG	1238
Qy	332	TyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrVal	351
Db	1239	TTTGACAAAGGAAGTCAACCCCTCGACCTTGCTCCAAATCGACTCTTTGTCAGTCCGTT	1298
Qy	352	AspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaVal	371
Db	1299	GATCCCAAGGAACCTCCCTTGTCACAGCGNACACTGCTCTTTCATCTCTTCTGTG	1358
Qy	372	AspTyrProValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThr	391
Db	1359	GATTATCCGGTTGAGAAGTCTCTCGTCTATGTTTCTGATGATGGTGTGCAATGCTTACG	1418
Qy	392	PheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLys	411
Db	1419	TTTGAAGCATGTCTGAACACTGGAATTTGCAAGAAATATGGGTTCCTTTTCAGCAAAAG	1478
Qy	412	PheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLys	431
Db	1479	TTTAAATATCGAGCTCGTGTCTCTGATGGTGTACTTCCACAGAGATAGACTACTGAAA	1538
Qy	432	AspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyArgGluGlu	451
Db	1539	GACAAGGTTGCTCTTCATTGTTAGGAGAGAGGGCGATGAAGAGAGAATACGAGAA	1598
Qy	452	PheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrp	471
Db	1599	TTCAGGTAAAGATCAATGCTCTGGTTGCAAAAGCCCAAAAGGTTCTCGAGGAAGATGG	1658
Qy	472	IleMetLysAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGlyMetIle	491
Db	1659	ACATGCAAGATGGAGGCCCTGGCCTGGAAACAACGTACCGCATCATCTCGAATGATT	1718
Qy	492	GlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeu	511

1719	Db	CAGGTATTCTCTGGCCAAAGTGGCGGTGCTGATGTCGGAAGGAATCAGATTGCTCGCCTG	1777
512	Qy	ValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMet	531
1779	Db	GTTTATGTCTCGAGAGAAAGAGCCAGGTTATAACCATCACAGAAGGCTGGTGCCATG	1838
532	Qy	AsnAlaLeuLeuArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAsp	551
1839	Db	AATGCACCTGCTCGCTCTCTGCTGCTTATCAAAATGCTGCATACCTATTATGAACATTGGAC	1898
552	Qy	CysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAsp	571
1899	Db	TGTGATCATTACATCAACAATACGAAGGCCATAAAGAGGCTATGTGTTTCATGATGAT	1958
572	Qy	ProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAsp	591
1959	Db	CTTTTGGTGGGAAGAAAGTGCTATGTATACAGTTCCTCAGAGGTTTGATGATTGAC	2018
592	Qy	ValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeu	611
2019	Db	AAAAATGATCGATACGCTACAGGAACGTTGTCTTTTTTGACATCAACATGAAGAGTTTG	2078
612	Qy	AspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeu	631
2079	Db	GACGGTATTCAAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGCACTG	2138
632	Qy	TyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAspCysCys	651
2139	Db	TATGTTTATGATGCTCTCTAAA---ACGAAGAAGCCACCATCAAGAACTTCGAACGTG	2195
652	Qy	Pro-----CysPheGlyArgGlyLysArgLys-----	660
2196	Db	CCCAAGTGGTCTCTCTCTGCTGTCGACGAGCAACAAGAAATAAAGAAGACTACAAAA	2255
660	Qy	-----	660
2256	Db	CCAAGACGGAGAGAAAGAAAGATTATTTTCAAGAAACGACGAAAACCCATCTCCTGCA	2315
661	Qy	HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGlu	680
2316	Db	TATGCTTTGGGTGAAATTGATGAAGTGCTCCA-----GGTGTGATATCGAGAAGGCC	2369
681	Qy	MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr	700
2370	Db	GGAATCGTAATCAACAGAAACTAGAGAAGAAATTTGGGAGCTTCTCTGTTTTTGTGCG	2429
701	Qy	SerThrLeuMetGluGluGlyGlyValProProSerSerProAlaAlaLeuLeuLys	720
2430	Db	TCAACACTCTTGAGAACGGAGGACCCCTGAAGAGCGCAGTCAGCTTCTCTCTGAAG	2489
721	Qy	GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu	740
2490	Db	GAAGCTATACATGTTATCAGCTGGGCTACGAGACAAAGCCGACTGGGGAAGAAGATT	2549
741	Qy	GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg	760
2550	Db	GGCTGGATTTCAGGATCGATCACAGAGGATATCTTGACTGGATTAAAGATGCATGCCAT	2609
761	Qy	GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle	780
2610	Db	GGCTGGCGGTCTATTTTACTGATCCCCGAAGCGGCTGCATTCAAAGGTTCTGCGCCTCTG	2669
781	Qy	AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe	800
2670	Db	AACCTTTCCGACCGTCTTCACCGAGTCTTCGCTGGGCCCTTGGTCCGTCGAAATTTTC	2722
801	Qy	PheSerArgHisSerProLeuLeu-TyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGln	820
2730	Db	TTTCAGNAANCAATGGNCACCTTTGGTACCGGATAC---GGCGCGCGGCTTAAATTCCTCGA	2788
820	Qy	uArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr	840
2787	Db	AAGTTTTCTTATATCAACTCACTGCTTTATCCCTGGACGTCCTTCTCTGCGCTTA	2840

QY 934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuLeuPro 953
 Db 3212 GACGAGGACGGGACACTTCGGAGCTGTACATGTTCAAGTGACGACGCTCTGTATCCCG 3271
 QY 954 ProThrThrLeuLeuLeuLeuValIleGlyValValAlaGlyLeuSerAspAlaIle 973
 Db 3272 CCACACCATCTGATCATCAACCTGGTGGCGTCTGCGGGCATCTCTTACGCCATC 3331
 QY 974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
 Db 3332 AACAGCGGATACAGTCGTTGGGGCCGCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTC 3391
 QY 994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
 Db 3392 ATCTGTCACCTGTACCGCTTCTCAAGGCGCTCATGGGCGAGCAGAACCGCACCGACC 3451
 QY 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
 Db 3452 ATCTGCTGCTGGGCCATCTGCTGGCGTCCATCTTCTCTGTGTGGGTTCGCATC 3511
 QY 1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db 3512 GACCCCTTCACCAACCGGCTCACTGGCCCGGATACCCAGACGCTGGCATCAACTGC 3568

RESULT 6

CL972423 3192 bp DNA linear GSS 21-SEP-2004
 LOCUS OsIFCC022918 Oryza sativa Express Library Oryza sativa (indica
 DEFINITION cultivar-group) genomic, genomic survey sequence.
 ACCESSION CL972423
 VERSION CL972423.1 GI:52399375
 KEYWORDS GSS.

ORGANISM

Oryza sativa (indica cultivar-group)
 Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoae; Oryza.
 1 (bases 1 to 3192)

REFERENCE

AUTHORS Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
 Wang, G. K. S., Deng, X. W. and Wang, J.
 An analysis of transcriptional regulation of the rice genome and
 its comparison to Arabidopsis
 Unpublished (2004)
 CONTACT: Chen Chen

JOURNAL

COMMENT Department of Bioinformatic
 Beijing Institute of Genomics
 Chinese Academy of Sciences, Beijing 101300, China
 Tel: 86-10-80481559
 Fax: 86-10-80488676
 Email: chenchen@genomics.org.cn
 Rice genomic sequence.
 Class: exon-trapped.

FEATURES

source
 1..3192
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="genomic DNA"
 /db_xref="taxon:39946"
 /clone_lib="Oryza sativa Express Library"
 /note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3192
 Score: 3907.00 Matches: 753
 Percent Similarity: 79.42% Conservative: 100
 Best Local Similarity: 70.11% Mismatches: 144
 Query Match: 69.17% Indels: 78
 DB: 9 Gaps: 14
 US-10-627-132-30 (1-1052) x CL972423 (1-3192)

QY 33 SerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeu 52
 Db 43 AGCGGCGAGGCGTGCAGATCTCGCGCAGCGCGTGGGCACGCGCGGCGGCGGCGGCG 102
 QY 53 PheValAlaCysAsnGluCysGlyPheProValCysArgProCysTrpGlyTrpGluArg 72
 Db 103 TTCCGCGCTTGGAGCGTCTCGCGCTTCCCGGTGTCGCCCTCTGTACAGTACGAGCGC 162
 QY 73 ArgGluGlyThrGlnAsnCysProGlnCysValysThrArgTyrLysArgLeuLysGlySer 92
 Db 163 AAGGATGCGACCCAGGCTTGGCCCGAGTCAAGACCAAGTACAAGGCCACAAAGGGAGC 222
 QY 93 ProArgValAlaGlyAspAspGluGlu---AspIleAspAspLeu---GluHisGlu 110
 Db 223 CCGGCGATCCGTTGGGAGGAGGCGGATGATCTGATGCTGATGATGCTCAGTCACTACAAC 282
 QY 111 PheAsnIleAspAspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGln 130
 Db 283 TACCCTGCATCTGGCAGTGGCGACCGAAGCAG-----AAG 318
 QY 131 IleThrGluAlaMetLeuHisGlyArgMetSerTyrArgGlyArgGlyProAspAspGlyAsp 150
 Db 319 ATTGCTGATAGGATCGCGAGTTGGCGCATGAATGCTGGGGTGGTGGAGATCTCGGCCCT 378
 QY 151 GlyAsnAsnThrProGlnIleProProIleIleThrGly--SerArgSerValProVal 170
 Db 379 -----CCCAAGTATGACAGTGGCGAGATCGGCTCACAATACT 417
 QY 170 erGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerLeuHisL 190
 Db 418 CGGAGAAATCCCTGGAGCTTCCCTCATCATATGATGCTCTTACCGAAACATT 477
 QY 190 ysArgIleHisProTyrProGlyProGlySerAlaLys-----T 205
 Db 478 GGCAAGCGTGTCCATTCCTTATGTGAACCATTCACCAATCCATCAAGGAGTCTCT 537
 QY 205 rpAspGluLysGluValSerTrpLysGluArgMetAspAspTrpLysSerLysGlnG 225
 Db 538 GGCAGCATTTGAAAT--GTTGCTGGAAAGAAAGATTGATGGCTGGAAACTGAAGCAG- 594
 QY 225 lyIleLeuGlyGlyGlyAlaAspPro----- 233
 Db 595 -----GACAAGGGAGCAATTCCTATGACCAATGGGACAGCATTCCTCCCTTCTGAAG 646
 QY 234 -----GluAspMetAspAla-----AspValAlaLeuA 243
 Db 647 CTCGGGAGTGTGTGATATCGATCGATCCACTGATTACAATATGGAAGATGCCTTACTGA 706
 QY 243 snAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLysValAsn 263
 Db 707 ATGATGAAACTCGCCAGCCTCTCTAGAAAAGTTCCCTTCTTCATCCAGATAAATC 766
 QY 263 roTyrArgMetValIleValValArgLeuValLeuAlaPhePheLeuArgTyrArgI 283
 Db 767 CTTACAGATGTGTCATGTTCTGCTGGTGTGTGTCTAAGCATTTTCTTCTACATCCGTA 826
 QY 283 leLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIle 303
 Db 827 TTACGATCTCTGCGCAATCGTATCGCTCTCGCTTTATCTGTTATATGTGAGATT 886
 QY 303 rpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGlu 323
 Db 887 GGTTCCTTCTGCTGGATATTGGATCAGTTCGGAAGTGGTTTCCAATCAACCGTAAA 946
 QY 323 hrTyrLeuAspArgSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerA 343
 Db 947 CCTACCTTGATAGGCTGGCATTTGAGGTATGACAGAAAGGAGGAGCCATCTCAGTTGGCTG 1006
 QY 343 laValAspLeuPheValSerThrValAspProLeuLysGluProProLeuValThrAla 363
 Db 1007 CTGTTGACATTTTGTTCAGTACAGTCGACCCCATGAGGAACCTCTCTTGTACTGCCA 1066
 QY 363 snThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrVal 383

```
Db 1067 ATACCGTGTCTGCATCTTCCTGCTGATTAACCCAGTGGACAGGTCTCTTCGTATGTAT 1126
Qy 383 erAepAepGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaA 403
Db 1127 CTGATGATGGTGTGCAATGCTGACTTTTGATGCAATGGCTGAGACTTCAGAGTTTGCTA 1186
Qy 403 rglYsTrpValProPheCysAllysPheGlyIleGluProArgAlaProGluPheTrp 423
Db 1187 GAAAGTGGGTTCCTATTCGTTAAAGAGTATTAACATTTGAGCCCAAGAGCTCCTGAGTGGTACT 1246
Qy 423 heSerLeuLysValAspTrpLeuLysAspLysValGlnProThrPheValGlnGluArgA 443
Db 1247 TCTCCAGAAATTCGATTACTTTGAAGACAAGTCCACCTTCATTTGTTAAAGACCGCC 1306
Qy 443 rgaLaMetLysArgGluTrpGluGluPheLysValArgIleAsnAlaLeuValAlaLysA 463
Db 1307 GTGCCATGAAGAGAGATATGAAGAAATTCAAAGTTAGGATAAATGGCCTTGTGCTAAGG 1366
Qy 463 laMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnA 483
Db 1367 CACAGAAGTCCCTGAGGAGGATGGATCATGCAAGATGGCACCACTGGCCAGGAACA 1426
Qy 483 snThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspT 503
Db 1427 ATACTAGGACCATCTCGGAATGATTCAAGTTTCTTGGTTCACAGTGGTGGCCTTGATA 1486
Qy 503 hrGluGlyAsnGluLeuProArgLeuValTrpValSerArgGluLysArgProGlyPheG 523
Db 1487 CCGAGGGTAATGAGCTTCCCGCGCTGGTCTACGTATCTCGTGAGAAGCGTCTCGGTTC 1546
Qy 523 lnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrA 543
Db 1547 AGCAACCAAGAGAGCTGGTGCCATGAATGCTGTGTTGCTGCTCAGCTGTGCTTACCA 1606
Qy 543 snAlaProPheMetLeuAsnLeuAspCysAspHisTrpIleAsnAsnSerLysAlaIleA 563
Db 1607 ATGCACAGTACATGTTGAAATCTTGATGTGATCACTACATCAACACAGCAAGGCTCTGC 1666
Qy 563 rglGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTrpValGlnP 583
Db 1667 GGGAGCTATGTGCTTCTTATGGACCCAAACCTAGGAAGGAGTGTCTGTTATGTTCAGT 1726
Qy 583 heProGlnArgPheAspGlyIleAspValHisAspArgTrpAlaasnArgAsnThrValP 603
Db 1727 TCCCAACAAGGTTGATGGTATGTAGTAATGATCGATATCGCAACAGGAACACTGTGT 1786
Qy 603 hePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTrpValGlyThrG 623
Db 1787 TTTTCGATATTAACTTGAGGGGTCTTGATGGCATCCAAGGACCAGTTTATGTGGGAAC 1846
Qy 623 lyCysValPheArgArgGlnAlaLeuTrpGlyTrpAsnProPro---LysGlyProLysA 642
Db 1847 GTTGTGTATTCAACAGAAGTCTCTATATGGTTATGAACCCCAATTAAGCAGAGAAGA 1906
Qy 642 rglProLysMetValThrCysAspCysCysProCysPheGlyArgLysLys----- 658
Db 1907 AGGGAAGTTTCTGTGCA-----TCACTTTGTGGGGGAGGAAGGAAGCAAGCAAGT 1957
Qy 659 -----ArgLysHisAlaLysAspGlyLeuPro- 667
Db 1958 CAAAGAAGAGAGTTCCGGACAAAAGAGTCAACACAGCAGTGGCAGTGTGTGCCAG 2017
Qy 668 -----GluGlyThrAlaAspMetGlyValAspSerAspLysG 680
Db 2018 TTTTCAAATCTTGAAGATATAGAGGAGGGTGTGAAAGGTGCTGGATTTCGATGATGAGAAT 2077
Qy 680 luMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValT 700
Db 2078 CACTTCTTATGTCTCAATAGACTTGGAGAAGAGATTTGGCCAGTCTGACGGCTTGTG 2137
Qy 700 hrSerThrLeuMetGluGluGlyGlyValProProSerSerProAlaAlaLeuLeuL 720
```

RESULT 7
AY110415
LOCUS

AY110415 3898 bp mRNA linear HTC 17-OCT-2002

```
2138 CTTCCACTCTCATGGAATATATGGTGGTGTCTCCTCAATCTGCACCCAGACATCTCTTTTGA 2197
Qy 720 ysgluAlaIleHisValIleSerCysGlyTrpGluAspLysThrAspTrpGlyLeuGluL 740
Db 2198 AAGAAGCTATCCATGATGAATGTTGGCTGATGAGGACAGACCGAATGGGAACTGAGA 2257
Qy 740 euGlyTrpIleTrpGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysA 760
Db 2258 TTGGGTGGATCTACGGTTCGGTGACAGAGATATTTCTACTGGATTCAAGATCATCGCG 2317
Qy 760 rglYTrpArgSerValTrpCysMetProLysArgAlaAlaPheLysGlySerAlaProI 780
Db 2318 GAGCTGGAGATCAATCTACTGTCATGCCAAGCGCCAGCTTTCAAGGGGTCTGCTCCTA 2377
Qy 780 leasnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleP 800
Db 2378 TCAATCTTTCCAGATCGCTTTAAACCAAGTGTCTCGGTGGGCACATTGGTCTCTGTGA 2437
Qy 800 hePheSerArgHisSerProLeuLeuTrpGlyTrpLysAsnGlyAsnLeuLysTrpLeuG 820
Db 2438 TTTTCAGTCGCCATTGTCCCATATGTTACGGCTAT---GGAGGACGCTTAAAGTTCITGG 2494
Qy 820 luArgPheAlaTrpIleAsnThrThrIleTrpProPheThrSerLeuProLeuLeuAlaT 840
Db 2495 AGAGATTTGGCTACATCAACACCACTATTTATTCATTTGACATCGATCCCGCTTCTCATAT 2554
Qy 840 yrcysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerT 860
Db 2555 ACTGTGTTTGGCTGCTATCTGTTGCTCACTGGAAAGTTCAATCCAGAGATTAGCA 2614
Qy 860 hrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuG 880
Db 2615 ACTTGTGATGATTTGGTTCATCTCTCTCTTCATTTCAATTTTGGCAGTGGTATCCTTG 2674
Qy 880 luMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValI 900
Db 2675 AGATGAGTGGAGTGGTGTGTGGCATCGATGAGTGTGGAGGAATGAACAGTTCTGGGTTA 2734
Qy 900 leGlyGlyValSerAlaHisLeuPheAlaValGlnGlyLeuLeuLysValLeuAlaG 920
Db 2735 TTGAGGTATATCTGCGCATCTTTTTCGGCTCTTCCAGGGTCTCTCTCAGGTGCTGCTG 2794
Qy 920 lyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAspGluPheA 940
Db 2795 GTATCCGACACCAATTTCACTGTCACTCAAGGCTCTCT---GATGAAGATGGTGACTTTG 2851
Qy 940 laGluLeuTrpAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleI 960
Db 2852 CTGAGCTCTACATGTTCAAGTGGACAAACGCTTCTCATCCCAACCGACACCATCTGATCA 2911
Qy 960 leasnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTrpGlnSerT 980
Db 2912 TTAACCTGGTGGGTGTGCTGGTATCTCATACGGCATCAACAGGGCTTACCAATCAT 2971
Qy 980 rglYProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTrpProP 1000
Db 2972 GGGACCGCTCTTTGGCAGAGCTCTTCTTGGCTCTTGGTGTGATTTGTCACATTGACCCCT 3031
Qy 1000 heLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSerI 1020
Db 3032 TCCTCAAGGGTCTTATGGGTGGCAAAACCGCATCCGACCATCGTGTGTTGTTGGGCAA 3091
Qy 1020 leLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgT 1040
Db 3092 TCCCTTCTGGCTTCGATCTTCTCATGTGTGGTTCGATCGATCGATCCATTCACCACCGAG 3151
Qy 1040 hrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3152 TCACCGGCCACAGATACCCCAACATGTGGATCACTCACTGC 3189
```



```

DEFINITION Zea mays CL1166_1 mRNA sequence.
ACCESSION AY110415
VERSION AY110415.1 GI:21214824
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S.,
        Harney, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences
JOURNAL Overgo Probes
REFERENCE Unpublished (2002)
AUTHORS 2 (bases 1 to 3898)
Coe, E.H.
DIRECT SUBMISSION
SUBMITTED (25-APR-2002) Maize Mapping Project, University of
MISSOURI, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
        these are publicly available from ZmDB and may be found by BLAST
        searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
        www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
        maize cDNA sequences is either Virginia Walbot, Stanford or Pat
        Schnable, Iowa State, then clones may be requested from ZmDB:
        www.zmdb.iastate.edu.
FEATURES
        Location/Qualifiers
            1..3898
               /organism="Zea mays"
               /mol_type="mRNA"
               /db_xref="MaizeDB:630057"
               /db_xref="taxon:4577"
               /clone_lib="Maize Mapping Project/DuPont Consensus
               Library"
               /note="this sequence is part of a project of EST
               assemblies resulting from the application of public
               contigs to seed Dupont contigs; this resource was
               assembled by Dupont as part of a collaboration for the
               overgo addressing of BACs in conjunction with the Maize
               Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3898
Score: 3821.50 Matches: 727
Percent Similarity: 75.42% Conservative: 120
Best Local Similarity: 64.74% Mismatches: 157
Query Match: 67.66% Indels: 119
DB: 3 Gaps: 17

US-10-627-132-30 (1-1052) x AY110415 (1-3898)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu120
DB 189 ATGGCGGCCAACAGCGGATGGTGGCGCTCGCAACACCGCAACGAGTTCGTCTATC248
QY 21 ArgGlyHisGluAsp-----ProLysProLeuArgAlaLeuSerGlyGlnVal36
DB 249 CGCCACGACGGCAGCGCGCGCGCGCTAAGCCACGAGAGTGCAGATGGCGAGTTC308
QY 37 CysGluileCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys56
DB 309 TGCCAGATTGGCGGCACTCTGTGGCGGTTTCAGCCACTGGTGATGTCTTTGTCCTGC368
QY 57 AsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGlyThr76
DB 369 AATGAGTGGCTTCCTCTGTCGCCGCCCTTCTATGATGATGACGACGAGGAGGGAAC428
QY 77 GlnAsnCysProGlnCysAlaThrArgTyrLysArgLeuLysGlySerProArgValAla96
DB 429 CAATGCTGCCCTTCAGTGCAGACTAGATACAGACAGACAGAAAGGTAGCCCTCGAGTTCAT488

```

```

QY 97 GlyAspAspAspGluGluAspIleAspIleAspLeuGluHisGluPheAsnIleAspGlu116
DB 489 CGTGATGAGGATGAGAAAGATGTTGATGACCTAGACAAATGAATTCACATAC-----539
QY 117 AsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeu136
DB 539 -----539
QY 137 HisGlyArgMetSerTyrGlyArgGlyPro-----AspAspGlyAsp150
DB 540 -----AAGCAAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG593
QY 151 GlyAsnAsnThrPro-----GlnIleProProIleIleThrGlySer164
DB 594 CTGTCTTCATCTGCTCGCCATGACCCACACCATCGATTCCACGCTGACAGCGGTCAA653
QY 165 ArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGly-----182
DB 654 CAG-----ATATCTGGAGAGATCCCTGATGATCCCTGACCGCTCAITCTATCCGC704
QY 183 GluValSerSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySer202
DB 705 AGTCCACATCGAGCTATGTTGATCCAGCGTCCCGAGTTCCTGTGAGGATTGTG-----758
QY 203 AlalysTrpAspGluLysGlu-----ValSerTrpLys214
DB 759 -----GACCCCTCGAGGACTTGAATTCCTATGGGCTTAATAGTGTGACGTGAAG809
QY 215 GluArgMetAspAspTrpLysSerLysGln-----224
DB 810 GAAAGAGTTGAGAGCTGGAGGGTTAAACAGGACNNNNNTATGTTCCAAAGTGACTAATAAA869
QY 225 -----GlyIleLeuGlyGlyAlaAspProGluAspMetAsp237
DB 870 TATCCAGAGGCTAGAGGAGGAGACATGAGGGAGCTGCTCAATGGAGAGATATGCAA929
QY 238 AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla257
DB 930 -----ATGTTGATGATGACGACGCTACCTTTGAGCGCGCATCGTGCAATTTCC977
QY 258 SerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhe277
DB 978 TCAACACGAGCTCAACCTTTACCGGTAGTATGATCTCCGCTTATCATCTCTGCTTC1037
QY 278 PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyIleTrpLeuValSer297
DB 1038 TTCTTCCAGTATCGTATCAGTCCATCGAGTGGTGTATGATGATGATGATGATGATCT1097
QY 298 IleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhe317
DB 1098 GTTATCTGCGAGGCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT1157
QY 318 ProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGlu337
DB 1158 CCAATCAACCGTGAGACATATCTGACAGGCTTCATTTGAGGTATGATGAGAGGAGAG1217
QY 338 ProSerLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluPro357
DB 1218 CCATCAGAGTGGCTCCCATGATGATGCTTCCTGTCAGTACAGTGGATCCATTGAGGAACCT1277
QY 358 ProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys377
DB 1278 CCATGATCAGACCAACACTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT1337
QY 378 ValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGlu397
DB 1338 GTGTCATGCTATGTTCTGATGATGGCTCAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT1397
QY 398 ThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArg417
DB 1398 ACCCGCGAATTTGCTAGAAAGTGGGTTCCCTTTTGTAGAGACCAATATTTGAACCAAGA1457
QY 418 AlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThr437

```

```
Db 1458 GCTCAGAAATTTACTTGTCTCNNNNNNTAGATTACCTGAAGGACAAAATTCACACCTTCA 1517
Qy 438 PheValGlnGluArgAlaMetIysArgGluTyrGluGluPheIysValArgIleAsn 457
Db 1518 TTTGTGAAGGAACGCGCAATGAAGAGGAGTATGAAGAAATTCAAAGTAAGAAATCAAT 1577
Qy 458 AlaLeuValAlaIysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThr 477
Db 1578 GCCCTTGTTCGAAGACAGAAAGTGCCTGAAGAGGGGTGGACATGGCTGATGGAACT 1637
Qy 478 ProTrpProGlyAenAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 497
Db 1638 GCTTGGCTGGCAATAACCTAGGACCATCTCGCATGATTCAGGTGTTCTTCTGGGCGAC 1697
Qy 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 517
Db 1698 AGTGGTGGCTCGACACTGATGGAATAAGATTAGATTACACAGCTCTGTCTCTCGGTGAA 1757
Qy 518 LysArgProGlyPheGlnHisIleLysIysAlaGlyAlaMetAsnAlaLeuIleArgVal 537
Db 1758 AAGAGACGAGCTTTCAGCATCAAGAAGAGCTGGTGCATGAATGCCTGATTCGTGTA 1817
Qy 538 SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn 557
Db 1818 TCTGCTGTCTGACAAATGGTGCCTATCTCTCAATGTGGATTGCGACCATTAATTCAAT 1877
Qy 558 AsnSerIysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
Db 1878 AGCAGCAAGCTCTAGAGAGCAATGTGCTTCATGATGGATCCGGCTTAGGAAGGAAA 1937
Qy 578 ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAla 597
Db 1938 ACTTGTATGTACAAATTTCCAGAGATTTGATGGCATTTGATGGCAGATCGATATGCT 1997
Qy 598 AsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
Db 1998 AATCGGAACATAGTCTTTTGATATCAACATGAAGGCTGGATGGCATTCAGGGTCCA 2057
Qy 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro--- 636
Db 2058 GTCTACGTGGGAACAGGATGCTGTTTCAATAGGCAGGCTTTGTATGGATACGATCCTGTT 2117
Qy 637 ---ProLysGlyProLysArgProLysMetValThrCysAspCys----- 650
Db 2118 TTGACTGAAGCTGATCTGGAGCTTAACATTTGTTAAGAGCTGCNNNNNNNNNNNNNN 2177
Qy 650 ----- 650
Db 2178 NNNNNNNNNNNNNNATGATAGTCAAAAGCCGTATTATTAAGAGAAACAGAAATCTTCA 2237
Qy 651 CysProCysPheGlyArgLysIysArgLysHisAlaLysAspGlyLeuProGluGlyThr 670
Db 2238 GCTCCCATCTTCAAC-----ATGGAAGACATCGAGGAGGGTATT 2276
Qy 671 AlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLys 690
Db 2277 GAA-----GGTTACAGAGATGAAGGTCAGTGTCTTATGTTCCAGAGGAAATTTGAGAAA 2330
Qy 691 ArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValPro 710
Db 2331 GCCTTGTGTCAGTCTCCAATCTTCATTCGATCCACCTTTATGACTCAAGGTGGCATACCA 2390
Qy 711 ProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 730
Db 2391 CCTTCAACAAACCCAGCTTCTCTACTAAAGGAAGCATTCATTCATCAGTTGTGGATAT 2450
Qy 731 GluAspLysThrAspTrpGlyLeuGluGlyTyrIleTyrGlySerIleThrGluAsp 750
Db 2451 GAGGACAAACTGAATGGGAAAAGAGATGGCTGGATCTATGTTTCAGTTACGGAGAT 2510
Qy 751 IleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLys 770
```

```
Db 2511 ATTCTGACTGGGTTTAAANNNNNNNNAGGGCTGGCAATCAATCTACTCATGATCCACCA 2570
Qy 771 ArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 790
Db 2571 CGACCTTGTTCGAAGGGTCTCGACCAATCAATCTTCTGATCGTCTTATCAGGTGCTC 2630
Qy 791 ArgTrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGly 810
Db 2631 CGTTGGGCTCTTGGTTCAGTGGAAATTCCTCTTAGCAGACATTCCTATCTATGATGGT 2690
Qy 811 TyrIysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrIleTyr 830
Db 2691 TAC----AATGACGCAATTCGAAGCTTTTGGAGAGGCTGGCTTACATCAACACATTTAT 2747
Qy 831 ProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThr 850
Db 2748 CCAATCACAATCTGTTCGGCTTATCGCTATTTGTGTCTCTGCTATCTCTCTTACC 2807
Qy 851 GlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPhe 870
Db 2808 AATAAATTTATCATCTCTGAGATTAGCAATATATGCTGGGATGTTCTTCAATCTCTTTC 2867
Qy 871 MetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGlu 890
Db 2868 GCCTCCANNNGCCACTGGTATATTCGAGCTTAGATGGAGTGGTGTGGCATTTGAAGAT 2927
Qy 891 TrpTrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaVal 910
Db 2928 TGTGGGAAATGAGCAGTTTTGGTATTGTTGGGACCTCTGCCCATCTCTTCGCAGTG 2987
Qy 911 ValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 930
Db 2988 TTCCAGGCTCTGCTGAAAGTGTGGCTGGGATGATACCAACTTCACAGTTACCTCAAAG 3047
Qy 931 AlaThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrLeu 950
Db 3048 GCATCT---GATGAGGATGGCAGCTTTGCTGAGCTATATGTTTCAAGTGGACAGTTG 3104
Qy 951 LeuIleProProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSer 970
Db 3105 CTCATTCTCGACCATCTTCTTGTTCATTAACTGCTGGATGGTGGCAGGAAATTCG 3164
Qy 971 AspAlaIleAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAla 990
Db 3165 TATGCCATTAACAGTGGCTACCAATCTCGGGTCCGCTCTTTGGAAAGCTTCTTCTCG 3224
Qy 991 PheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArg 1010
Db 3225 ATCTGGGTGNTCTCCCATCTGTACACGTTCCTCAAGGGTCTCATGGGAAGGACGACCG 3284
Qy 1011 ThrProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuTrp 1030
Db 3285 ACACCAACATCTGTCATGTCGTCCATCTCTTCTTGCATCTATCTTCTCTCTGCTGG 3344
Qy 1031 ValArgIleAspPropheIleValArgThr---LysGlyProAspValArgGlnCysGly 1049
Db 3345 GTGAAGATCGATCCTTTTCATCTCCCGCACAGAAAGCTGCTGCTTGGGGCAATGTGGC 3404
Qy 1050 IleAsnCys 1052
Db 3405 GTCAACTGC 3413
RESULT 8
CNS0A1ML 3911 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSJPGH52A06 of Hormone Treated Callus of strain col-0 of
ACCESSION BX832166
VERSION BX832166.1 GI:42459101
KEYWORDS HTC; SLT; cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS

1 (bases 1 to 3911)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

TITLE

Unpublished

JOURNAL

2 (bases 1 to 3911)

REFERENCE

Genoscope.

AUTHORS

Direct Submission

JOURNAL

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

source

1..3911

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTPGH562A06"

/tissue_type="Hormone Treated Callus"

/plasmid="pCMVSPORT_6"

1..3911

/gene="At5g64740"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3911
Score: 3719.00 Matches: 711
Percent Similarity: 76.57% Conservative: 132
Best Local Similarity: 64.58% Mismatches: 193
Query Match: 65.85% Indels: 65
DB: 3 Gaps: 17

US-10-627-132-30 (1-1052) x CNS0A1ML (1-3911)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnAtcAsnGluLeuValLeuLeu 20
Db 430 ATGAACACCGGTGGTGGTAAATCCCGGTTCTTCACACAGGAATGAGTTTGTCTCAT 489
QY 21 ArgGlyHisGluAspPro-----LysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 490 AATCGCATGAGAATGCCGGAATAAGATCATCAGTCCAGAGCTGAGTGGACAGACATGTCAA 549
QY 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 550 ATCTGCAGAGATGAGATCAATGACTGTGTATGAGAGAACCGTTTGTGCGATGTAAACGAA 609
QY 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGlyThrGlnAsn 78
Db 610 TGTGCATTCCTCTGTGTAGACCTTGTCTATGAGTACGAAAGACGAGAGGCAATCAAGCT 669
QY 79 CysProGlnCysLeuThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 670 TGTCCACAGTGCAAAACCGTTTCAACGTTCTTAAGGAAGTCCAAAGAGTTGAAGGTGAT 729

QY 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 730 GAAGAGGAGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 781 -----AATGGGATTTGGATTTGATCAGGTTTCTGAAGGATGATGATCAATCTCT 825
QY 139 ArgMetSerTyrGlyArgGlyProAspGlyAspGlyAsnAsnThrPro----- 155
Db 826 CGTCGCAACTCGGTTTCCCAACTCTGATTTGGAT-----TCAGCTCCACTGGCTCT 879
QY 156 GlnIleProProIleIleThrGlySerArgSerValProValSerGlyGlu----- 172
Db 880 CAGATTCCATTGCTGACTTACGCCACAGGACGTTGAGATTCTTCTCTGATAGACATGCT 939
QY 173 -----PheProIleThrAsnGlyTyrGlyHisGly---GluValSerSer-SerLeuHi 189
Db 940 CTTATTGTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 996
QY 189 slys-----ArgIleHisProTyrProValSerGluProGlySerAlaLysTrpAs 206
Db 997 TGACCCGACCGTGGCTGCACATCCAGCGCTATGCTACCTCAGAAGATCTTCGCGTTTA 1056
QY 206 pGluLysLysGluValSerTrpLysGluArgMetAspTrpLysSerLysGlnGly-- 225
Db 1057 TGGTTATGGAAGTGTGCTTGGAAAGATCGGATGAGGAGTGAAGAGAGAGAGAGAGAG 1116
QY 226 -----IleLeuGlyGlyGlyAlaAspPro-----GluAspMetAspAlaAs 239
Db 1117 GAACTTCAGGTTGTTAGGCATGGAAGGAGATCCTGATTTTGAAGATGTTGATGATGCTGA 1176
QY 239 pValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSe 259
Db 1177 TTTTCCATGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1236
QY 259 rLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePheLe 279
Db 1237 CAAGATAAATCCTTACCGGATGTTAAATGTGTGCTACGTCTTGTGATCTTGTGCTCTCT 1296
QY 279 uArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleI 299
Db 1297 TCATCTACCGTATCTTTCACCCCGTCAAAGATGCTATGCTTTGTGGCTTATTTCTGTTAT 1356
QY 299 eCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProI 319
Db 1357 ATGTGAGATATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 1416
QY 319 eAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGlyGluProSe 339
Db 1417 CGAGCGAGAAACGTACTTGGACCGACTCTCATTAAGATATGAGAAGAAGGAGAAACCGTC 1476
QY 339 rLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProle 359
Db 1477 GGGACTATCCCCCTGGGATGATTTTGTGTAGTACAGTGGATCCATTGAAAGAGCCCTCCGCT 1536
QY 359 uValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSe 379
Db 1537 TATTACTGCAATATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1596
QY 379 rCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAl 399
Db 1597 TTGTTAGTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1656
QY 399 aGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaPr 419
Db 1657 TGAATTGCAAGGAAATGGGTTCTCTTCTGCAAGAAATATTGTTGAGCCCTCGTCTCC 1716
QY 419 oGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheVa 439
Db 1717 CGAATGTTATTTCTGCCATTAATAGGACTCTTGAAGATAAAGTTTCATCCCGCATTTGT 1776
QY 439 lGlnGluArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLe 459

1777 TAGGAGCGCGGCGCATGAAGAGAGATTAAGAGAAATTCAAAAGTAAGATCAATGCTTT 1836
459 uValAlaLysAlaMetLysValProAlaGluGlyTrrPleMetLysAspGlyThrProTr 479
1837 AGTAGCAACAGCAGAGAAAGTGCCTGAAGATGGGTGGACTATGCAAGACGGTACACCTTG 1896
479 pProGlyAsnAsnThrArgAspHisProGlyMetIleGlnVal-PheLeuGlyHisSerG 499
1897 GCCCGGTAATAGTGTGGAGATCATCTCGCATGAATCAGGTCTTTCTTGGAAAGTGACC 1956
499 lyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysA 519
1957 GTGTTCTGTGTGTCGAAACCAACAGGTGCCTCGATTAGTTTACGTTTCTCGTGAGAAGA 2016
519 rgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerA 539
2017 GACCGGATTTGATCACCATAGAAGGCTGGAGCTATGAATTCCTCGATACGAGTCTCTG 2076
539 laValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnS 559
2077 GGGTTCATCAATAGCTCTTACCTTCTGAATGTCGATTGGATGCATACATCAACAATA 2136
559 erLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValC 579
2137 GCAAGCTCTTAGAGAAGCAATGTGTTTCATGATGGATCCTCAGTCAGGAAGAAATCT 2196
579 ySyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnA 599
2197 GTTATGTTTCAGTCTCCCTCAAGAGTTTCGATGGATTTGATAGGCACGATCGATCTCAAATC 2256
599 rgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValT 619
2257 GCAATGTTGTGTTCTTTGATATCAATATGAAGGTTTGGATGGGCTACAGGGCTATAT 2316
619 yrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysG 639
2317 ACGTCGGTACAGTTGTGTTTCAGGAGCGAAGCGCTTTACGGATTTGATGCACCGAAGA 2376
639 lyProLysArgProLysMetValThrCysAspCysCysPro-----CysP 654
2377 AGAAGAGGGGCCACAGTAAG---ACATGCAATTTGCTGCCCAAAATGGTGTCTCTATGTT 2433
654 heGly---ArgLysLysArgLys----- 660
2434 TTGGTTCAAGAAAGAAATCGTAAGCAAGACAGTGGCTGGATGAAGAGAGAGATA 2493
661 -----HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetG 674
2494 GGGAGCGCTCAAGCAGATCCACGCAATTAGAAATATCGAAGAGGCGCGCTCACTAAAG 2553
674 lyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyG 694
2554 GTTCTAAGCTAGAACAGTCAACCGAGGCAATGCAAAATGAAGTTGGAGAAGAAATTTGGGC 2613
694 lnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSers 714
2614 AGTCTCTGTATTTGTTGCATCTGCGGTATGGAGATGGTGGATGGCTAGAAACGCA 2673
714 exProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysT 734
2674 GCCCGGCTGTCTGCTTAAAGAGCCATCCAAAGTCATTAGTTGCGGATATGAAGATAAA 2733
734 hrAspTrpGlyLeuGluLeuGlyTrrPleTyrGlySerIleThrGluAspIleLeuThrG 754
2734 CTGAATGGGAAAGAGATGGGTGGATCTATGTTCTCTGTTACCGAAGATATTTCTACGG 2793
754 lyPheLysMetHisCysArgGlyTrrArgSerValTyrCysMetProLysArgAlaAlaP 774
2794 GTTTTAAAGATGATCTCTATGTTGGAGATCTGTTTATTTGATACCAAAAGTAGCGGCTT 2853
774 heLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrrPAlaL 794

2854 TCAAAGGATCAGCTCCAATCAATCTTTTCGATCGTCTCCATCAAGTTCTTCGATGGGCGC 2913
794 euGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnG 814
2914 TTGGGTCGGTTGAGATTTTCTGAGTAGGCAATGCTCTATTTGTTATGGTTAT---GGAG 2970
814 lyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrIleTyrProPheThrS 834
2971 GTGGGTTGAATGGCTTGAGCGGTTGCTCTACATTAACCTCTGTGGTTTACCCGTGGACCT 3030
834 erLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheI 854
3031 CTCTACCGCTCATCGTTTACTGTTCTCTCCCTCCCATCTGCTCTCTCACTGGAAATTC 3090
854 leMetProSerIleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIleP 874
3091 TCGTTCGAGATAGTAACTATGCGAGTATCTCTTCATGGCGCTCTTCTCGTCGATG 3150
874 heAlaThrGlyIleLeuGluMetArgTrrPserGlyValSerIleGluGluTrrPArgA 894
3151 CAATAACGGGTATCTCGAGATGCAATGGGCAAGTTGGGATCGATGTTGGTGGAGAA 3210
894 snGluGlnPheTrp-VallIleGlyValSerAlaHisLeuPheAlaValGlnGly 913
3211 ACGAACAGTTTTCGGGTCATTGGAGGTTTCTCGCATCTGTTGCTCTCTTCCAAAGGT 3270
914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
3271 CTCTCAAGGTTCTGCTGGTGCAGCACTAATCTTCAATGGACTTCACTTCTCATCCCT 3327
934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIlePro 953
3328 ---GATGATGGAGGTTCTCTGACCTTTTACTCTTCAATGGACTTCACTTCTCATCCCT 3384
954 ProThrThrLeuLeuIleIleAsnValIleGlyValAlaGlyIleSerAspAlaIle 973
3385 CCAATGACTCTACTCATATAAAGTCATTTGGAGTCATAGTCGGAGTCTCTGATGCCATC 3444
974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpVal 993
3445 AGCAATGGATACGACTCGTGGGACCGCTTTTTCGGAAGACTGTTCTTTGCACTTGGGTC 3504
994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
3505 ATCATTCACTTTTACCGCTTCTTAAAGTTTGTCTGGGAAAACAAGATAGATGCCAAC 3564
1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
3565 ATTATTGCTGCTGCTCCATCCCGCTCGGCTCTTACACTTCTTTGGGCGCGGTT 3624
1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
3625 AATCCGTTTGTGGCGAAA---GGCGGCTCTATTCTCGAGATCTGTGTTAGACTGC 3678
RESULT 9
AV112236 3728 bp mRNA linear HTC 17-OCT-2002
LOCUS
DEFINITION Zea mays CL1160_1 mRNA sequence.
ACCESSION AV112236
VERSION AV112236.1 GI:21216826
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3728)
AUTHORS Hainey, C. P., Dolan, M., Miao, G. H., Vogel, J. M., Whitsitt, M. S.,
Arthur, L. W., Hanafey, M., Morgante, M. and Tingey, S. V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3728)

AUTHORS Coe, E.H.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
 COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES
 Location/Qualifiers
 1..3728
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="MaizeDB:630049"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3728
 Score: 3633.50 Matches: 703
 Percent Similarity: 75.38% Conservatives: 99
 Best Local Similarity: 66.07% Mismatches: 211
 Query Match: 64.33% Indels: 51
 DB: 3 Gaps: 12

US-10-627-132-30 (1-1052) x AY112236 (1-3728)

```

QY 37 CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 292 TCCAGATCTGGCGGATCGCGTGGGACACTACGGCGGAGGAGAGCGTCTTCCCGCTGC 351

QY 57 AsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGlyThr 76
   :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 352 GACGTCGCGGTTCGGGTGTCGCCCTGCTAGAGTACGAGCGCAAGCGGCACA 411

QY 77 GlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAla 96
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 412 CAAGCGTGNNNNAGTGNNNNCAAGTACAGCGCCACAGGGGAGTCCAGCGATCCA 471

QY 97 GlyAspAsp---AspGluAspIleAspAspLeuGluHisGluPheAsnIle----- 113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 472 GGGGAGGAGGAGGACGATCTGATGCGGATGATGCTAGC---GACTTCAACTACCGTGA 528

QY 114 ---AspAspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThr 132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 529 TCTGCAATGACACAGACAGAGAGATTGCTGACAGGATCGGCGATGCGGCGATGAAT 588

QY 133 GluAlaMetLeuHis-----GlyArgMetSerTyrGlyArgGlyProAspAspGlyAsp 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 589 GCTNNNNNCAGCGGGGATGTTGGCGCCCAAGTATGACAGTGGTGGATCGGGCTTACC 648

QY 151 GlyAsnAsnThrProGlnIlePro---ProIleIleThrGlySerArgSerValProVal 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 649 AAGTACGACAGTGGTGAGATCTCGGGGATACATCCNNNNNNNNNNNNNNNNNNAT 708

QY 170 SerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeuHis 189
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 709 TCGGAGAAATCCCTGGTCTCCCTGACCATATATGATGCTCTCTACTCGGGAACATT 768

QY 190 LysArgIleHisProTyrProVal-----SerGluProGlySerAlaLysTrp 205
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 769 GCGAGCGCGGCCCATTTCCCTATATGAATCATCATCAATCCGTCGAGGGAATCTCT 828
  
```

```

QY 206 AspGluLysLysGluValSerTrpLysGluArgMetAspAspTrpLysSerLysGln---- 224
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 829 GGTAGCGTGGGAATGTTGCTCGAAAGAGAGGGTGTGATGGCTGGAAATGAAGCAGAC 888

QY 225 ---GlyIleLeu-----GlyGlyGlyAlaAspPro 233
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 889 AAGGAGACAAATTCCTCATGCAAGCAACGATTCCTCTCTGAGGGCGCGGGGTGT 948

QY 234 GluAspMetAspAla-----AspValAlaLeuAsnAspGluAla 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 949 GGTGATATTGATGATCACTCAATGATGGAAGATGCTTATTAACGATGAACACT 1008

QY 247 ArgGlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMet 266
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1009 CGCCAGCCTCTATCTAGGAAGTTCCTCTCCAGGATAAATCATACAGGATG 1068

QY 267 ValIleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisPro 286
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1069 GTCATGTGTACGATTGATTGTTCTAGCATCTTCTGCACTACCGGATCAAAATCCT 1128

QY 287 ValProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIle 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1129 GTGCGTAATGCATACCCACTGTGGCTTCTATCTGTTATATGTAGATCTGGTTGCTCT 1188

QY 307 SerTrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAsp 326
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1189 TCCTGGATATTGGATCAGTTTCCAAGTGGTTTCCAATCAACCGGAGACTTACCTTGTAT 1248

QY 327 ArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeu 346
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1249 AGACTCGCATTAAGGTATGACCGGAAGGTAGCCATCTCAGTTGGCTGCTGTGACANN 1308

QY 347 PheValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeu 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1309 NNGTCACTGCTGCGACCAATGAAGGAGCTCTCTTGTCTACCTGCCAATACCTGTCTA 1368

QY 367 SerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspGly 386
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1369 TCCATTCGCTGAGTATCTCTGAGTAAAGGTCTCTGCTATGTATGTATCTGATGATGA 1428

QY 387 AlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpVal 406
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1429 GCTGTATGTGACATTTGATGACACTGAGACTTACAGACTTTCAGATAAATAGGGTGT 1488

QY 407 ProPheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLys 426
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1489 CCAATTTTAAGAGTACCAATGAACTAGAGCTCTCTGATGCTACTTCTCCAGAAA 1548

QY 427 ValAspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLys 446
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1549 ATTGATTACTTGAAGGACAAAAGTGCACCTTCAATTTGTTAAAGACCGCGGNNNNNN 1608

QY 447 ArgGluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysVal 466
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1609 NNNNNNNNNNNNNNNNNNNCAATAGGGTAATAGGCTTGTGTGTAAGGACACNNNGTC 1668

QY 467 ProAlaGluGlyTyrPheMetLysAspGlyThrProTyrProGlyValAsnThrArgAsp 486
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1669 CCGAGGAGGATGGATCATCAAGATGCGACCATGCGCCAGGAAACAATACAGGAGAC 1728

QY 487 HisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluLysAsn 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1729 CATCTGGAATGATTTCAGGTTTCTCGTGGTGGTGGTCTTGTGATCTGAGGAGTAAT 1788

QY 507 GluLeuProArgLeuValTyrSerArgGluLysArgProGlyPheGluHisLysLys 526
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1789 GAGTACCCCTTGTGTTCTATGTTCTCGTNNNNNNCGTCTCTGGATCCAGCATCACAG 1848

QY 527 LysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPhe 546
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1849 AAGCTGTGGCATGAATGCTCTTGTCCGCTCTCAGCTGTGCTTACCAATGGAACAATAC 1908

QY 547 MetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMet 566
  
```

```
Db 1909 ATGTTGAATCTTGATTTGATCACTACATCAACAGTAGTAAAGGCTCTCAGGAGAGCTATG 1968
Qy 567 CysPheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArg 586
Db 1969 TGCCTTCCTTATGGATCTCAACCTAGGAGGAGTGTCTGCTATGTTCAGTTTCCCCAGAG 2028
Qy 587 PheAspGlyLeuAspValHisAspArgTyrAlaAsnArgAsnThrValPheAspIle 606
Db 2029 TTCATGGTATTGATAGGAATGATCGATATGCCAACAGGAACACCGTGNNNNCGATATT 2088
Qy 607 AsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPhe 626
Db 2089 AACCTTGAGAGCTCTTGATGGCATCCAAGGACCACTTTATGTGGGCACTGGCTGTGTTTC 2148
Qy 627 ArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetVal 646
Db 2149 AACAGAACAGCTCTATATGTTATGAGNNNNNAATTAAGCAAAAGAGGGTGGTTTCTTG 2208
Qy 647 ThrCysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLysAsp- 664
Db 2209 TCATCACTATGTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2268
Qy 665 -----GlyLeuProGlu 668
Db 2269 NNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTAGAGAG 2328
Qy 669 GlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPhe 688
Db 2329 GGAGTTGAAGCGCTGGATTTGATGATGAGAAATCACTCTTATGTCTCAAAATGAGCTG 2388
Qy 689 GluLysArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGly 708
Db 2389 GAGAAGAGATTGGCCAAATCGCAGCTTTTGTGGTCCACTCTGATGGAATNNNNNNNN 2448
Qy 709 ValProProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCys 728
Db 2449 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTATCAATGTTGT 2508
Qy 729 GlyTyrGluAspLysThrAspTyrGlyLeuGluLeuGlyTyrIleTyrGlySerIleThr 748
Db 2509 GGCTACGAGGACAAAGATTGAATGGGGAACCTGAGATTTGGGTGGATCTATGTTCTGTGACG 2568
Qy 749 GluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMet 768
Db 2569 GAAGATATTCATCTGGGTTCAAGATGACGACGAGGCTGGCGGTGCAATCTACTGCAATG 2628
Qy 769 ProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGln 788
Db 2629 CCTAAGCGCGCGCTTCAAGGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACCCAG 2688
Qy 789 ValLeuArgTyrAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeu 808
Db 2689 GTGCTCCGGTGGGCTCTCGGTTCAAGTGGGAAATCTCTTTTCAGCGCGCATTTGNNNNNTATGG 2748
Qy 809 TyrGlyTyrLysAsnGlyAsnLeuLysTyrLeuGluArgPheAlaTyrIleAsnThrThr 828
Db 2749 TACGGGTAC---GGAGGAGCGCTGAAGTCTTGGAGAGATTCGCTTACATCAACACCCAC 2805
Qy 829 IleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeu 848
Db 2806 ATCTACCCGCTCAGCTCCCTCCGCTCCCTCAATTTACTGTATCTCTGCTGCCATCTGCCTG 2865
Qy 849 LeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAla 868
Db 2866 CTCACGGGGAAGTTTCATCCACAGAGATCAGCAACTTGGCTAGTATCTGTTTCATCTCT 2925
Qy 869 LeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSerIle 888
Db 2926 CTCCTTCATCTCGATCTTCCCAAGGATCTCTGGAGATGAGTGGAGGGCGCTGGGCATC 2985
Qy 889 GluGluTyrTyrArgAsnGluGlnPheThrValIleGlyGlyValSerAlaHisLeuPhe 908
:::|
```

```
2986 GACGAGTGGTGGAGAACGAGCAGTCTTGGGTTCATCGGAGGCATCTCCGCCACCTCTTC 3045
Qy 909 AlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr 928
Db 3046 GCCGTCTTCCAGGCGCTCTCTCAAGGTGTTCCGGCATCGACACCACTTCACCGTCACC 3105
Qy 929 SerLysAlaIleThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTyrThr 948
Db 3106 TCCAAGGCGCTCG--GATGAAGACGCGACTTCGCGGAGCTGTACATGTTTCAAGTGGACG 3162
Qy 949 ThrLeuLeuIleProProThrThrLeuLeuIleAsnValIleGlyValValAlaGly 968
Db 3163 ACATCTCTGTATCCCGCCACCACTCTGATCATCAACCTGTGTGGCGTGTGTTCGCCGC 3222
Qy 969 IleSerAspAlaIleAsnAsnGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhe 988
Db 3223 ATCTCTACGCCATCAACAGCGGTACCACTGTGGGTCCGCTCTTCGGCAGACTCTTC 3282
Qy 989 PheAlaPheThrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGln 1008
Db 3283 TTCGCTTCTGGGTGATCGTTACCTGTACCGCTTCTCAAGGCTCTCATGGGTGGCAG 3342
Qy 1009 AsnArgThrProThrValValValIleTyrSerIleLeuLeuAlaSerIlePheSerLeu 1028
Db 3343 AACCGCACCCCGACCACTCGTGTGTCTGGGCGATCTCTGGCGTGCATCTTCTCTCTG 3402
Qy 1049 GlyIleAsnCys 1052
Db 3463 GGCATCAACTGC 3474
RESULT 10
LOCUS AY110079 3696 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL1164_1 mRNA sequence.
ACCESSION AY110079
VERSION AY110079.1 GI:21214162
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3696)
AUTHORS Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1..3696
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630055"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
```



```
Db 2165 GCCTATATGGTGTATGAGCTCCAGTCAAGNNNNNNNGCCAGGCTTCTTCTCTCGCTT 2224
Qy 650 CysCys-----ProCysPheGlyArgLysLysArg 659
Db 2225 TGTNNNNNAAGNNNNNGACGTCAAATCTAAGAAGAGCTCGGAAAAGAAAGTCCACAT 2284
Qy 660 LysHisAlaLysAspGlyLeuPro-----GluGlyThrAla 671
Db 2285 AGACACGACGACAGCTTCTGTACCAAGTATTTAATCTCGAAGATATAGAGGAAGGATTTGAA 2344
Qy 672 AspMetGlyValAspSerAspLysGluMetSerHisMetAsnPheGluLysArg 691
Db 2345 GGTTCCTCAGTTGATGATGAGAAATCGCTGATTAATGCTCAATGAGCTTGGAGAAAGA 2404
Qy 692 PheGlyGlnSerAlaPheValThrSerThrLeuMetGluGluGlyValProPro 711
Db 2405 TTTGGCCAGTCCAGTGNNNNGTAGCCTCTACTCTGATGGAATATGTTGGTGTGTCCACAA 2464
Qy 712 SerSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyThrGlu 731
Db 2465 TCTGCAACTCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2524
Qy 732 AspLysThrAspTTPGlyLeuGluGlyTrpIleTyrGlySerIleThrGluAspIle 751
Db 2525 GACAAACTGACTGGGAGACTGAGATTGGGTGGATCTATGTTCTGTACAGAGACATT 2584
Qy 752 LeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArg 771
Db 2585 CTCACCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2644
Qy 772 AlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg 791
Db 2645 CCAGCTTTCAAGGAGATCGCTCCATCACTTTCGGATCGTTTGAATCAAGTGTCTGG 2704
Qy 792 TrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyr 811
Db 2705 TGGGCTCTTGGTTCATTTGAATTTCTTTCAGCAGGCATTTCCCATATGATGATGGCTAT 2764
Qy 812 LysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrTrpIleTyrPro 831
Db 2765 ---GGAGCCCGCTTAAATTTCTGGAGAGATTTGCTTATATCAACACACAAATTTATCCA 2821
Qy 832 PheThrSerLeuProLeuLeuAlaTyrCysThrIleProAlaValCysLeuLeuThrGly 851
Db 2822 CTCACATCAATCCCGCTCTCTGCTACTGCTATATTGCCAGCAGTTTGTCTCTCACTGG 2881
Qy 852 LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMet 871
Db 2882 AAGTTTCATATCCCAAGATTTAGTAACCTAGAGAGTGTGGTTTATATCGCTCTTTATC 2941
Qy 872 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluTrp 891
Db 2942 TCAATCTTTGCCACTGGTATCTCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3001
Qy 892 TrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaVal 911
Db 3002 TGGAGGAACGACGACTTCTGGGTCTATTTCTGGGTATTTCTGGGCATTTATTTGCCGCTTC 3061
Qy 912 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 931
Db 3062 CAGGCTCTCTGAAGGTGCTTCTGGTATCGACAGAGCTTCACCTCTAAGGCC 3121
Qy 932 ThrGlyAspGluAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeu 951
Db 3122 ACT--GACCAAGAGAGTGAATTTGCCGAGCTCTACATGTTCAAGTGGCAACGCTTCTG 3178
Qy 952 IleProProThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAsp 971
Db 3179 ATCCACCAACCACTATTTTGAATTCATCAACCTGGTGGGTGCTGCTGGCAATTTCTTAC 3238
Qy 972 AlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPhe 991
```

```
Db 3239 GCAATCAATAGCGGTTTACCAGTTCATGGGACCTCTTTTTCGGGAAGCTCTTCTTTCGCTT 3298
Qy 992 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
Db 3299 TGGGTGATTGTCCACTGTACCTCCCTTCCCTCAAGGGCTCATGGGAAGCAGACCGCAGC 3358
Qy 1012 ProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1031
Db 3359 CCGACCATGTGCTGTCTGGGTATCTCTCTGCGTCCGATCTTTTCCCTGATGTGGTT 3418
Qy 1032 ArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn 1051
Db 3419 CGTATCGATCCATTCACCCCGGCTCACTGGTCTCTGATATCGGAAATGTGCATCAAC 3478
Qy 1052 Cys 1052
Db 3479 TGC 3481

RESULT 11
LOCUS AY104236 Zea mays PC0121439 mRNA 2872 bp linear HTC 16-OCT-2002
DEFINITION Zea mays PC0121439 mRNA sequence.
ACCESSION AY104236
VERSION AY104236.1 GI:21207314
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2872)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Zea Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2872)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1..2872
location/Qualifiers
1..2872
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:637064"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
```

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2872
Score: 3350.50 Matches: 617
Percent Similarity: 84.84% Conservative: 88
Best Local Similarity: 74.25% Mismatches: 96
Query Match: 59.32% Indels: 30
DB: 8 Gaps: 8

US-10-627-132-30 (1-1052) x AY104236 (1-2872)

QY 246 AlaArgGlnProLeuSerArgIysValSerIleAlaSerSerIysValAsnProTyrArg 265
DB 13 GCACGGCTACCTCTAAGTCGCAAGTCCGATATCTCCAAACAGCTTAACCTTTATCGG 72
QY 266 MetValIleValValLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHis 285
DB 73 ATCGTGATTTGTTCTCGCGCTTATCATCTTCTTTCTTTCAATATCGTAACTCAT 132
QY 286 ProValProAspAlaIleGlyLeuTyrLeuValSerIleIleCysGluIleTyrPheAla 305
DB 133 CCAGTGGAAAGATGCTTATGGTTGTGGCTTGTATCTGTATTTGTGAAGTTGGTTGGC 192
QY 306 IleSerTyrIleLeuAspGlnPheProIysTyrPheProIleAspArgGluTyrIleLeu 325
DB 193 TTGTCTTGGCTTCTAGATCAGTTCCAAAGTGTATCTTATCAACCGTGAACCTTACCTC 252
QY 326 AspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAla-ValAs 345
DB 253 GATAGACTTGCAATGATATGATAGGAGGGTGGCCATCCAGTTGGGCTCCCAATCGA 312
QY 345 pLeuPheValSerThrValAspProLeuIysGluProProLeuValThrAlaAsnThrVa 365
DB 313 TGTCTTTTGTAGTACAGTGGATCCACTTAAGGAACCTCTCTAAATTACTGGCACTGT 372
QY 365 IleuSerIleLeuAlaValAspTyrProValAspIysValSerCysTyrValSerAspAs 385
DB 373 CCGTGTCCATTTCTGTGTGGATTAACCTGTGCAAGATATCATGTATGTATCTTCATGA 432
QY 385 pGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgIleThr 405
DB 433 CGGTTCAGCTATGTGACTTTTGAAGCGCTATCTGAACCGCAGAGTTTGGCAAGAAATG 492
QY 405 pValProPheCysIysPheGlyIleGluProArgAlaProGluPheTyrPheSerIle 425
DB 493 GGTTCCTTTTTCAGAAACACAAATATTGAACCTAGGGCTCCAGAGTTTACTTTTGCTCG 552
QY 425 uIysValAspTyrLeuIysAspIysValGlnProThrPheValGlnGluArgAlaMe 445
DB 553 AAAGATAGATTACTTAAGCAACAAATACAACTTCTTTTGTGAAGAAAGCGGGCTAT 612
QY 445 tIysArgGluTyrGluGluPheIysValArgIleAsnAlaLeuValAlaIleAlaMetIly 465
DB 613 GAAGAGGGAGTGTGAAGATTCAAGTACGGATCGATGCCCTTGTGCAAAAGCGCAAAA 672
QY 465 sValProAlaGluGlyTyrIleMetIysAspGlyThrProTyrProGlyIysAsnThrAr 485
DB 673 AATACCTGAGGAGGGCTGACCATGGCTGATGGCACTCCYTGCGCTTGGGAATAACCTAG 732
QY 485 gAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGl 505
DB 733 AGATCATCCMGGAATGATCCAAGTATTTCTTGGGCCACAGTGGTGGCTTGACACGGATGG 792
QY 505 yAsnGluLeuProArgLeuValTyrValSerArgGluIysArgProGlyPheGlnHisHi 525
DB 793 GAATGAGTTGCCACGGCTTGTATGTATTTCTCGTGAAGAGCGCCAGGCTTCCAGCACC 852
QY 525 sIysIysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaPr 545
DB 853 CAAGAAGGCTGTGCCAATGCTTTGTATTCGGGTATCAGGTGTCTTGACGAAATGGTGC 912
QY 545 oPheMetLeuAsnLeuAspCysAspHisTyrIleAsnSerIysAlaIleArgGluAl 565
DB 913 TTATCTTCTTAATGTGGATGTGTGATCACTACTTCAATAGCAGCAAAAGCTCTTAGAGAGC 972
QY 565 sMetCysPheLeuMetAspProGlnValGlyArgIysValCysTyrValGlnPheProGl 585
DB 973 TATGTGTTTCATGATGGATCCAGCTAGGAAGGAAACCTTCTGTATGTTCAGTTTCCACA 1032
QY 585 nArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAs 605
DB 1033 AAGATTTGATGTATAGACTTCATGATCATATGCAATGCAACCGGAACATTTGCTTTTGA 1092

QY 605 pIleAsnMetIysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysVa 625
DB 1093 TATTAATATGAAGGCTAGATGGCATTCAGGACCTGTTTATGTGGGAACAGGATGCTG 1152
QY 625 lPheArgArgGlnAlaLeuTyrGlyTyrAsnPro-----ProIysGlyProIysArgPr 643
DB 1153 TTTCAATAGGAGGCTTGTATGGCTATGATCTCTGTATTGACAGAGCTGATTGGAGCC 1212
QY 643 oIysMetValThrCysAspCysCysProCysPheGlyArgIysLysArg----- 659
DB 1213 TAAACATTATCAATAAAGTTGCTGTGGC-----GGAAGAAAAAGAGACAGAGCTA 1266
QY 660 -----LysHisAlaIysAspGlyLeuPro----- 667
DB 1267 TATTGATTCCAAAAACCGTGATATGAAGAGAACAGAACTTCCTGGCTCCCATCTTCAACAT 1326
QY 668 -----GluGlyThrAlaAspMetGlyValAspSerAspIysGluMetLeuMe 683
DB 1327 GGAAGATATAGNAGAGGATTTGAA-----GGTTACGAGGATGAAGAGTCACTGCTTAT 1380
QY 683 tSerHisMetAsnPheGluIysArgPheGlyGlnSerAlaAlaPheValThrSerThrLe 703
DB 1381 GTCTCAGAAGAGCTTTGGAGAAACGCTTTGGCCAGTCTCCAATTTTATTGCATCCACCTT 1440
QY 703 uMetGluGluGlyGlyValProProSerSerSerProAlaAlaLeuLeuIysGluAlaI 723
DB 1441 TATGACTCAAGGTGCATACCCCTTCAACAAACCCAGGTTCCCTGCTAAAGAGAACTAT 1500
QY 723 eHisValIleSerCysGlyTyrGluAspIysThrAspTyrGlyLeuGluLeuGlyTyrI 743
DB 1501 ACATGCTCATTTAGTTGTGATATGAGGATAAACAGAAATGGGGGAAGAGATCGGATGGAT 1560
QY 743 eTyrGlySerIleThrGluAspIleLeuThrGlyPheIysMetHisCysArgGlyTyrPar 763
DB 1561 ATATGGCTCTGTACTGAAGATATTTTAACTGGTTTCAAGATGCATGCAAGAGTTGGAT 1620
QY 763 sSerValTyrCysMetProIysArgAlaAlaPheIysGlySerAlaProIleAsnLeuSe 783
DB 1621 ATCCATCTACTGCATGCCACTTCGGCTTGGCTTCAAGGGTTCCTGCTCCGATTAATCTTTC 1680
QY 783 rAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIlePhePheSerAr 803
DB 1681 TGATCGTCTCAACCAAGTGTACGCTGGCTCTTGGTTCAAGTTGAAATTTACTTAGCAG 1740
QY 803 gHisSerProLeuLeuTyrGlyTyrIysAsnGlyAsnLeuIysTyrLeuGluArgPheAl 823
DB 1741 ACATGCTCTATCTGGTATGGTTAC-----AATGGAAGGCTAAAGCTTCTGGAGAGACTGC 1797
QY 823 aTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLe 843
DB 1798 ATACATCAACACCATTTGTTATCCAAATTCATCTATCCACTATAGTATAGCATCTAGCTCCT 1857
QY 843 uProAlaValCysLeuLeuThrGlyIysPheIleMetProSerIleSerThrPheAlaSe 863
DB 1858 TCCTGCTATCTGTTTACTCACCAAAATTTATTTATCTCGGATAGCAATTTATGCTGG 1917
QY 863 rLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTr 883
DB 1918 GCGCTTCTTCATCTGCTCTTTTGTCTTCATCTTCGCCACTGGTATTTGGAGCTTCGATG 1977
QY 883 pSerGlyValSerIleGluGluTyrTyrArgAsnGluGlnPheThrPValIleGlyIysVa 903
DB 1978 GAGTGGTGTGGCATTTAGGATTTGGTGAGAAATGAGCAGATTTTGGGTCAITTTGGTGGCAC 2037
QY 903 lSerAlaHisLeuPheAlaValGlnGlyLeuLeuIysValLeuAlaGlyIleAspTh 923
DB 2038 CTCTGCACATCTCTTGTGTTCCAAAGGTCTCTTAAAGTGTCTAGCAGGATCGACAC 2097
QY 923 rAsnPheThrValThrSerIysAlaThrGlyAspGluAspAspGluPheAlaGluLeuTy 943
DB 2098 AAACCTTCACGCTCAGATCAAAAGGCAACC-----GATGATGATGGTGTATTTGCTGAGCTGTA 2154
QY 943 rAlaPheIysTyrThrThrLeuLeuIleProProThrThrLeuLeuIleIleAsnValI 963

```

Db      2155  TGTGTTCAAGTGGACAACTCTTCTGATCCCCCCACCACTGTGCTGTGATTAACCTGGT 2214
Qy      963   eGlyValValAlaGlyLeuSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLe 983
Db      2215  TGGTATAGTGGCTGGAGTGTGATGCTATCAACAGTGGCTACCAATCATGGGTCCACT 2274
Qy      983   uPheGlyValLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysG1 1003
Db      2275  ATTGGGAGCTGTTCTTTGCAATCTGGTGATCTCCCACTCTACCTTTCTTGAAGGG 2334
Qy      1003  YLeuMetGlyArgGlnAsnArgTrpProThrValValValIleTrpSerIleLeuLeuAl 1023
Db      2335  TCTCATGGGAAGCAGAACCGCACACCGACCATCGTCAATTGTTGGTCCGCTCTTCTGC 2394
Qy      1023  aSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr---LysG1 1042
Db      2395  TTCCATATTCCTCGTCTGGTGGTGAAGATCGACCCCTTCATATCCCTACCCAGAGGC 2454
Qy      1042  yProAspValArgGlnCysGlyIleAsnCys 1052
Db      2455  TCTTCCGCTGGCAGTGTGTGTAACCTGC 2485

```

RESULT 12

CD726831

LOCUS

DEFINITION EST027 Cucurbita pepo testa subtracted cDNA Cucurbita pepo cDNA
clone CES similar to Cellulose synthase, mRNA sequence.

ACCESSION CD726831

VERSION CD726831.1

KEYWORDS GI:32277678

SOURCE EST.

ORGANISM

Cucurbita pepo

Cucurbita pepo

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.

1 (bases 1 to 1874)

Bezold, T.N., Mathews, D., Loy, J.B. and Minocha, S.C.

Molecular analysis of the hull-less seed trait in pumpkin:

Expression profiles of cell wall related genes during development

Unpublished (2003)

Contact: Subhash Minocha, Dennis Mathews, Brent Loy, Todd Bezold

Dr. Minocha

University of New Hampshire

Rudman Hall, Durham, NH 03824, USA

Tel: 603 862 3840

Fax: 603 862 3784

Email: sminocha@cisunix.unh.edu

Degenerate primers and Tag were used to amplify cDNA for TOPO TA

(Invitrogen, Carlsbad, CA) cloning. Sequencing was performed three

times using the Dyanamic ET Terminator Sequencing kit (Amersham

Pharmacia Biotech Inc, Piscataway, NJ). Sequences were identified

by NCBI BLAST(X).

Location/Qualifiers

1..1874

/organism="Cucurbita pepo"

/mol_type="mRNA"

/db_xref="taxon:3663"

/clone="CES"

/dev_stages="20 days post-anthesis"

/clone_lib="Cucurbita pepo testa subtracted cDNA"

/note="Organ: Testa; Total RNA was isolated from 20 day

post-anthesis testa tissue and used in a subtraction

hybridization procedure as according to the Clontech

PCR-Select cDNA Subtraction kit (PT1117-1) (Clontech, Palo

Alto, CA).

ORIGIN

Alignment Scores:

Pred. No.: 6,998-255 Length: 1874

Score: 2525.00 Matches: 467

Percent Similarity: 83.57% Conservative: 57

Best Local Similarity: 74.48% Mismatches: 77

Query Match:

DB: 44.71%

Indels: 26

Gaps: 4

US-10-627-132-30 (1-1052) x CD726831 (1-1874)

```

Qy      309  IleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeu 328
Db      2   GTGCTGGACCACTCCCTAAATGGTTACCTGTCTGCAGGGACACATTTATTATGACAGATTA 61
Qy      329  SerLeuArgTyrGluArgGlyGluProSerLeuLeuSerAlaValAspLeuPheVal 348
Db      62  TCAGCCAGGTTTGAAGAGAGGGAGAACCTTCCAGCTTGTCTGTGGATTTTGTGTG 121
Qy      349  SerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIle 368
Db      122  AGTACCGTTGATCCGTGGAAGAACCTTCTTAATCACTCGGATACCGTTCATTC 181
Qy      369  LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 388
Db      182  CTTGCTGTGACTATCTCGTAAAGTCTCTCTGCTATGTCGATGACGGTGCAGCT 241
Qy      389  MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPhe 408
Db      242  ATGCTCACTTGAATCTCTAGTTGAAACAGCTGACTTTGCAGGAAATGGTTCATTC 301
Qy      409  CysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAsp 428
Db      302  TGCMAAAATCTCCATTGAACCTCGAGCTCTGAGTTTACTTCTCCGAGAAATAGAC 361
Qy      429  TyrLeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGlu 448
Db      362  TATTTGAAAGATAAGTACAACCTTCTTTGTAAGGAACGCTAGAGCTATGAAGAGAGCT 421
Qy      449  TyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAla 468
Db      422  TATGAAGAGTTTAAAGTACGAGTCAATGCTTGTGGCAAGGCGCAAAACACCTGAT 481
Qy      469  GluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 488
Db      482  GAAGCTGTGTCATGCAAGATGGAACAGCTTGGCCAGGAACCAATCCACGTGATCACCT 541
Qy      489  GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeu 508
Db      542  GGAATGATTCAGGTTTTCTTGGAAATACCGGCCCATGACGTAGAGGGGAAATGAACCT 601
Qy      509  ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysValAla 528
Db      602  CCTCGACTAGTCTACGTTCTCCAGAGAGAGGCGCTGTTACCAGCATCACTCAAAAGCT 661
Qy      529  GlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeu 548
Db      662  GGTGCTGAAATGCTCTGGTAAGAGATCTGCACTCTCCCAATGCCCTTCATCTCT 721
Qy      549  AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPhe 568
Db      722  AACCTTGACTGTGATCACTATGTTAAACAATAGTCAAGCTATACGTAGGCAATGTGTTTC 781
Qy      569  LeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAsp 588
Db      782  TTGATGAGACCCACCAAGTTGGCCGAGATGTATGCTATGTTTCAGTTTCTCCAGAGTTT 841
Qy      589  GlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMet 608
Db      842  GGCATTGATCGAAGTATCGATATGCCAATCGCAACACGCTTCTTCGATGTGAACATG 901
Qy      609  LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg 628
Db      902  AAGGATTGGATGGCATCCAGGAGCAAGTTTATGTTGGTACAGGTTGTGTTTCAATAGA 961
Qy      629  GlnAlaLeuTyrGlyTyrAsnProPolysGlyProLysArgProLysMetValThr--- 647
Db      962  CAGGACCTTTATGGCTACGGACCTCTTCCATGCTAGCTTATCTAAGACATCTCTCATCA 1021

```


600 -----GCTGAAGCAGAACGGGGGGCGCCCGCGAGTTTCGACCAACACCGCTGGCTCTT 653
Db
199 ---GluProGlySerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArgMet 217
Qy
654 CGAGACCAAGGGCCCTACGGCTACGGCAACAGCAATCTGGCCCGAGGACGACGGCGT--- 710
Db
218 AspAspTrpLysSerLysGlnGlyIleLeuGlyGlyAlaAspProGluAspMetAsp 237
Qy
711 -----GGCGGGGCA---CCCCAAGGAGCTGAT 734
Db
238 AlaAspValAlaLeuAenAspGluAlaArg-GlnProLeuSerArgLysValSerIleAl 257
Qy
735 -----GAGCAAGCCATGCGCGCGCTGACCGCGCAAGCTCCCGATCCA 776
Db
257 aSerSerLysValAenProTrpArgMetValIleValValLeuValValLeuAlaPh 277
Qy
777 GCGCGCGGTGATACGCCGTACAGCTGCTGCTGATCCGGCTGGTGGCGCTGGGGTT 836
Db
277 ePheLeuArgTrpArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSe 297
Qy
837 GTTCTGATGGGGCATCAAGACCAAGACGAGGACCCATCTGGCTGTGGGGGATGTC 896
Db
297 rIleIleCysGlnIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPh 317
Qy
897 CATCGTGGCAGCTGTGTTGCGCTTGTGCTGGTGTGCGACCAAGCTGCGCCAAAGCTGTG 956
Db
317 eProIleAspArgGluThrTrpLeuAspArgLeuSerLeuArgTrpGlu-----ArgG1 335
Qy
957 CCCCATCAACCGCGCCACCGACCTGAGCGTGTCAAGGACCAAGTTTCGAGACGCCACGCC 1016
Db
335 uGlyGluPro-----SerLeuLeuSerAlaValAspLeuPheValSerThrValAs 352
Qy
1017 GAGCAACCCCGCCAGCGCATCGGACCTCCCGGGATCGACATCTCGTCTCCACCGCGCA 1076
Db
352 pProLeuLysGluProProLeuValThrAlaAenThrValLeuSerIleLeuAlaValAs 372
Qy
1077 CCGGAGAAGAGCCCGTGTGGTGACGGCCAAACACCATCTCTCCATCTCGCGCGCGCA 1136
Db
372 pTyrProValAspLysValSerCysTrpValSerAspAspGlyAlaSerMetLeuThrPh 392
Qy
1137 CTACCCCGTCGACAAGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1196
Db
392 eGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPh 412
Qy
1197 CGAGGCCATGGGAGCGCGCGAGCTTCGCCAACCTGTGGGTGCCATCTTCGCCGAAGCA 1256
Db
412 eGlyIleGluProArgAlaProGluPheTrpPheSerLeuLysValAspTrpLeuLysAs 432
Qy
1257 CGAGATCGAGCGGAGGAAACCCCGACAGCTACTTCAACCTCAAGAGGGACCCGTTCAAGAA 1316
Db
432 pLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTrpGluLuph 452
Qy
1317 CAAGGTGAAGGGCGCATCTGTCAAGGACAGCGCGCGGTGAAGCGGGAGTACGACGATT 1376
Db
452 eLysValArgIleAenAlaLeu----- 459
Qy
1377 CAAGTCCCGCTCAATGGCTTCCCGACGCCATCCCGCGCCGCTCCGACGCGTACCACGC 1436
Db
459 ----- 459
Qy
1437 CCGCGAGGAGATCCAGGCCATGAACCTCGACGGGGAGAGATGAAGCGCGCGGCGACGA 1496
Db
460 ----ValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPr 478
Qy
1497 GCAGCAGCTGGAGCGCGATAAAGATCCCAAGGCGAGCTGG---ATGGCGGACGGCACGCA 1553
Db
478 otPrProGly-----AenAenThrArgAspHisProGlyMe 490
Qy
1554 CTGGCGGGGACGTGGCTGACGGCGTCCGCGGAGCACGCGAGGGGACACACGCGGGAT 1613
Db
490 tIleGlnValPheLeu-----GlyHisSerGlyGlyHis----- 501
Qy
1614 CATACAGGTGATGCTAAAGCCTCCAGTCCAGTCCAGTCCAGCAGCAGCGCGCGGACATGGA 1673
Db

502 -----AspThrGluGly-----AenGluLeuProArgLeuValTyrValSerAr 516
Qy
1674 AAAGAGGGTGATTTGCGGGCGTGACACACCGCGCTGCCGATGCTGGTGTACGTGTGCG 1733
Db
516 gGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAenAlaLeuIleAr 536
Qy
1734 GGAGAAGCGCGCGGGTACGACCAACAAGAGAGGGCGGCGCATGAACCGCTGGTGGCG 1793
Db
536 qValSerAlaValLeuThrAenAlaProPheMetLeuAenLeuAspCysAspHisTrpI1 556
Qy
1794 GCGGTGCGCGATCATGTCCAACGGCGGTTCATCTCACTGACTGCGACCACTACGT 1853
Db
556 eAenAenSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyAr 576
Qy
1854 GTACACTCCAAAGCGTTCGGGAGGGGATGTCTTCATGATGAC---CGCGCGCGCA 1910
Db
576 gLysValCysTrpValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTy 596
Qy
1911 CCGCTCTGTCTAGTCCAGTTCCGACGCTTCGAGGCGATCGACCCCTCCGACCGCTA 1970
Db
596 rAlaAenArgAenThrValPhePheAspIleAenMetLysGlyLeuAspGlyIleGlnG1 616
Qy
1971 CGCAACCAACAACCGCTTCTTCGACGTCACATCGCGCCCTCGACGGCTTCCAGGG 2030
Db
616 yProValTrpValGlyThrGlyCysValPheArgArgGlnAlaLeuTrpGlyTrpAenPr 636
Qy
2031 CCGCTCTATGTCCGAACTGGCTGCTTTCGCGCGCATCGCGCTCTACGCTTTCGACCC 2090
Db
636 oProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyAr 656
Qy
2091 GCCCAGGTCCAAAGACCAACACCCCGCTGTCTCTGC-----TGCTGCCACG 2138
Db
656 qLysLysArgLysHisAlaLysAspGlyLeuProGluGlyThrAlaAsp---MetGlyVa 675
Qy
2139 TCGCGCGCGCAAGAATCCAGCCCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2198
Db
675 lAspSerAspLysGluMetLeuMetSerHisMetAenPheGluLysArgPheGlyGlnSe 695
Qy
2199 CCGCATGACATGACGCGCGCATGAACATGGCGAGCTTCCCAAGAAGTTCGGCACTC 2258
Db
695 rAlaAlaPheValThrSerThrLeuMetGluGlu----- 706
Qy
2259 GAGCTTCTGATGACTCCATCCCGTGGCGGAGTTCCAGGACGCGCGCTGGCGGACCA 2318
Db
707 -----GlyGlyValProPro-----SerSerSerPr 715
Qy
2319 CCGGTGCGTGAAGAACCGCGCGCGCGCGCGCGCGCTGACGATCCCGCGGGAGACGCTGA 2378
Db
715 oAlaAlaLeuLysGluAlaIleHisValIleSerCysGlyTrpGluAspLysThrAs 735
Qy
2379 CGCTGCGATCGTGGCGGCGCATCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2438
Db
735 pTrpGlyLeuGluLeuGlyTrpLleTrpGlySerIleThrGluAspIleLeuThrGlyPh 755
Qy
2439 GTGGGGGACGAGGGTGGGTTCGATATACGGTTCGTGACGGAGGACGTGGTCCACGGGTA 2498
Db
755 eLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLy 775
Qy
2499 CCGATGCACAAACCGGGGGTGAAGTCCGTCTACTGCGTCAACCGCGCGCGCTTCG 2558
Db
775 sGlySerAlaProIleAenLeuSerAspArgLeuAenGlnValLeuArgTrpAlaLeuG1 795
Qy
2559 CGGCACCGCCCATCAACCTCAGCAGCGCTCCACGAGTCTCTCCGATGGGCGCACTGG 2618
Db
795 ySerValGluIlePhePheSerArgHisSerProLeuLeuTrpGlyTrpLys----- 812
Qy
2619 CTCCGTGAGATCTTCTTCTCACCAACACGCACTCTTCGCTTCTCCTCAAGATGAAGT 2678
Db
812 ----- 812
Qy
2679 GCCTAGCTTGCTATTTCGATCGATTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2738
Db


```

QY 813 ----AsnGly-----AsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrTh 828
Db 2739 AACAAACAATATCCATTGAGGTTCAGGCGATCCAGCGATCTTACCTCAACCTGG 2798
QY 828 rLetyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLe 848
Db 2799 CATCTACCCCTTCACTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2858
QY 848 uLeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPheIleAl 868
Db 2859 CTCTCGGGGCGATTTCATGTCAGACGCTCAACGTCACCTCTCTCTCTCTCTCTCT 2918
QY 868 aleuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIl 888
Db 2919 CATCACCATCAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2978
QY 888 eGluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPh 908
Db 2979 GGAGGAGTGTGGCGGACGAGCAGTCTTGCTCTCTCTCTCTCTCTCTCTCTCTCT 3038
QY 908 eAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValTh 928
Db 3039 CGCGCTCTCTGCGGGGCTTCTCAAGGTCTATCGCGGATCGAGATCTCTTCAACCTC 3098
QY 928 rSerLysAlaThrGlyAspGlu---AspAspGluPheAlaGluLeuTyrAlaPheLysTr 947
Db 3099 CTCCAGCAGCTCGCGGACGAGCTGACGACGAGTTCGCGGAGCTCTACGCGTCAAGTG 3158
QY 947 pThrThrLeuLeuLeuProThrThrThrLeuLeuIleIleAsnValIleGlyValValAl 967
Db 3159 GAGTCTCTCATGATACCCCGCTTACCATCATATGATCATACCTGTCGCCATCGCGT 3218
QY 967 aGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLe 987
Db 3219 CGGCTTCAGCGGACCATCTACAGCACCATCCCGAGTGGAGCAAGCTTCTCGTGGAGT 3278
QY 987 uPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyAr 1007
Db 3279 CTCTCTCAGCTGTGGTCTCTCGCGCATCTCTACCCCTTCGCCAAGGGCTCATGGGCG 3338
QY 1007 gGlnAsnArgThrProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSe 1027
Db 3339 GAGGGGAGGACACCCACCATCTGTACGTCTGGTCTGGCTCTGGCTCTGGCTCACCATCTC 3398
QY 1027 rLeuLeuTrpValArgIleAspPro 1035
Db 3399 CTCTCTCTGATTGCAATCAAGCT 3423

RESULT 14
CL964957
LOCUS
DEFINITION
CL964957 Oryza sativa Expressed Sequence Tag (EST) cDNA, clone 1035, from a rice genomic library.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS.
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 2595)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
JOURNAL
COMMENT
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559

```

```

FEATURES
    source
        1..2595
            /organism="Oryza sativa (indica cultivar-group)"
            /mol_type="genomic DNA"
            /db_xref="taxon:3946"
            /clone_lib="Oryza sativa Expressed Library"
            /note="Oryza sativa exon trapped genomic sequences"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2,21e-213 Length: 2595
Score: 2134.50 Matches: 418
Percent Similarity: 63.5% Conservative: 127
Best Local Similarity: 48.7% Mismatches: 261
Query Match: 37.7% Indels: 51
DB: 9 Gaps: 13

US-10-627-132-30 (1-1052) x CL964957 (1-2595)

QY 232 AspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSer 251
Db 43 GATCTGAGCCTGCATCTTTCAGAAAGTGGTGTGATGACCATGACATACCTGAGCCTCTCAGC 102
QY 252 ArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleValValArg 271
Db 103 AGCAGGCTGTCGTACCTGCGGCGAGCTCAACATGTACCGCGCGCGCTCGCGCTCCGG 162
QY 272 LeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIle 291
Db 163 CTCGTCTCTCATCGCGCGCTTCTTCGCTACCGCGTCAACCGCGCGCTCGCGCGCCAC 222
QY 292 GlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAsp 311
Db 223 GCGCTGTGGGTGACCTCGCTCGCGTGTGAGTCTGCGCGCGCTGATGGTGTGATCGCC 282
QY 312 GlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArg 331
Db 293 CAGCTCCCGAGAGCTCTCCCGCGCCACCGCGCTGACGTACCTCGACAGGCTGCGCTCAGG 342
QY 332 TyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrVal 351
Db 343 TACGAGAAGCGCGGAGCGCTGCGGCTGCGGCGCTGCGGCGTGGACGTGTCTGCTCGCGCGG 402
QY 352 AspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaVal 371
Db 403 GACGCGCGAGGAGCGCGCTGCGCGACGCGAACACCGTCTGCTGCTGCTGCTGCTGCGCG 462
QY 372 AspTyrProValAspLysValSerCysTyrValSerAspAspGlyValAspMetLeuThr 391
Db 463 GACTACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522
QY 392 PheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLys 411
Db 523 TTCGAGTCTCTTTCAGCGCGCGCTGCGGAGCGGTGGATCCCTCTCTGCGCGCGCG 582
QY 412 PheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLys 431
Db 583 CACGCGCGGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642
QY 432 AspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluThrGluGlu 451
Db 643 GACGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702
QY 452 PheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrp 471
Db 703 TTCAAGGTGAGGATGAACCACTCTCGCGCGAGGCGCGCAAGGTGCGCGAGGAGATGG 762
QY 472 IleMetLysAspGlyThrProThrProGlyAsnAsnThrArgAspHisProGlyMetIle 491

```

```

Fax: 86-10-8048676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

```

```

Location/Qualifiers
    1..2595

```

Db 763 ATCATGTCCGACGCGCGTGGCGGGGAACAACCTCGCGGGATCAACCTCGCCATGATT 822
Qy 492 GlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeu 511
Db 823 CAGGTGCTCTGGGTCAACCGCGGACGCGACGTCGACGCGCGGAGCTCCCGCGCTC 882
Qy 512 ValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMet 531
Db 883 TTCTACGTGTCCGGGAGAGACGACCGGTTTCGGGACACACGCAAGCGCGGCCCATG 942
Qy 532 AsnAlaLeuLeuArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAsp 551
Db 943 AACGCGTGTCTACGGGTATCCGCGTGTCTTACCAATCGCGCTACGTCCTCAACCTGGAC 1002
Qy 552 CysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAsp 571
Db 1003 TGCACCACTCGTCAACAACAGACGCGCTGAGAGGCGATGTCTTTATGATGGAC 1062
Qy 572 ProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAsp 591
Db 1063 CCGTTCGCGGGAACAGACGTGTCTTCGTCAAGTTCGCACTGCGCGACAGCGGC----- 1116
Qy 592 ValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeu 611
Db 1117 -----GCGCGCGCGCGCGGCGACTCCGTCTTCTCGATATCGAGATCGAATGCTC 1170
Qy 612 AspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeu 631
Db 1171 GATGCGATCCAGGCGCGGTGTACGTGCGTTTCGCGTTGCTGCTTTCAGCGGAGCGCTG 1230
Qy 632 TyrGlyTyrAsnPro-----ProLysGlyProLysArgProLysMetValThr 647
Db 1231 TACGGTTCGAGCCCGCGCGCGCTGACGACGCGGACGACATGGATACGCGCGCGGAC 1290
Qy 648 CysAspCysProCysPheGlyArgLysLysArgLysHisAla----- 662
Db 1291 TGGAGGAGATGTCTGCTTCGGGAGGGGAGAGGATGAACGCGATGAGGAGGTCCATG 1350
Qy 663 -----LysAspGlyLeuProGluGlyThrAlaAspMetGly 674
Db 1351 TCCGCGGTGTCTGCTGACTCGGAGGATGACTCCGACGACAGGAGGAGGAGGCG 1410
Qy 675 ValAspSerAspLysGluMetSerHisMetAsnPheGluLysArgPheGlyGln 694
Db 1411 CGGGGAGAGAGAGAGCTGCGCGCTACCGCGCGCGCTGGAGCGCACTTCGCGCAG 1470
Qy 695 SerAlaAlaPheValThrSerThrLeuMetGlu-----GluGly 707
Db 1471 TCTCCCGGTTTCATCGCTCGCGTTCGAGGAGCAAGGCGCGGCGGCGGCGACGCT 1530
Qy 708 GlyValProProSerSer-----SerProAla-----AlaLeuLeuLysGluAlaIleHisVal 725
Db 1531 GGTCTGCTGACGCGACCGCTGCGCGACGCGATCTCTCAAGAGGCGCATCATGTC 1590
Qy 726 IleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTyrIleTyrGly 745
Db 1591 GTCACTGCGGCTTCAGAGAGGAGAGAGATGGGCAAGAGATTGGATGATGTACGCG 1650
Qy 746 SerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerVal 765
Db 1651 GGC-----GGTGTCCGACGCGGTTCCGGATGCACGCGCGGCTGGTGTGCGCG 1701
Qy 766 TyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArg 785
Db 1702 TACTGCTGCGCGGAGCGCGGTTCAGGCGGTACGCGCGCGCGCGCGCGGACGTC 1761
Qy 786 LeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhePheSer---ArgHis 804
Db 1762 CTCGGGGGGCGCTCGCGGCGCGGTGGCGCATGGGATCTCTGAGCGCGCGGAC 1821
Qy 805 SerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyr 824
Db 1822 AGCCCCGCTCGGCGCGCGCGCTCGCG-----CTCTGACGCGGCTGGGCTAC 1872

Qy 825 IleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuPro 844
Db 1873 GTGGCGCGCGCGCTACCGCTCGCTCCCTCCCGCTCACCGCTACTGCGCGCTCCG 1932
Qy 845 AlaValCysLeuLeuThrGlyLysPheIleMetProSer---IleSerThrPheAlaSer 863
Db 1933 GCGGTCTGCTCTCTACCGGCAAGTCCACGTTCCCGGACGACGTGAGCTACTACGACGC 1992
Qy 864 LeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTyr 883
Db 1993 GTCTCTCTCATCTGCTCTTCTCCGTGGCGCGTCCGCTCGCTCGAGCTGAGGTGG 2052
Qy 884 SerGlyValSerIleGluGluTyrTrpArgAsnGluGlnPheTrpValIleGlyVal 903
Db 2053 AGCGCGTGCATCGCGGCGTGTGGGAGAGAGAGCTCTGGATGGTCACCGCCACA 2112
Qy 904 SerAlaHisLeuPheAlaValValGlnGlyLeuLysValLeuAlaGlyIleAspThr 923
Db 2113 TCGCGAGCGCTCGCGCGCTGTTCAGGGCATCTCAGCGCTGACCGGATCGACGTC 2172
Qy 924 AsnPheThrValThrSerLysAlaThr----- 932
Db 2173 GCCTTCTCCAGGAGCGCGAGCTTCGCCACCCCAAGCGACCGCGCGGAGAAACGATGAC 2232
Qy 933 GlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIle 952
Db 2233 GGGAGAGAGAGCGACATTTGGATCGGATACGATCGGTGGTGGACGACCTGTGTGTG 2292
Qy 953 ProProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAla 972
Db 2293 GCGCGAGCGCTGGTGGTGGGACCTCGCGGGGTGTAGCGCGGTGGCGGTACGCGG 2352
Qy 973 IleAsnAsnGly---TyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 991
Db 2353 GTGACCAACGCGTACTACCACTCGTGGGCGCGCTCGCGCGAGCTGGCGCTCGCGCGG 2412
Qy 992 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
Db 2413 TGGTGTGGCGGACCTTCAGGATTCCTCGGGGGTCTCTGGCGCGCGAGACCGGCGG 2472
Qy 1012 ---ProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrp 1030
Db 2473 CCGCGGACCATCGCGGTCTCTGTCGTGTCTCTCGTCCGTGGGTCTGCTCTCTGG 2532
Qy 1031 ValArgIleAspProPheIleValArgThrLysGlyProAspValArgGln 1047
Db 2533 GTCCACCGCGCTCTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2583

RESULT 15

CL974686

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

3048 bp DNA linear GSS 21-SEP-2004
Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.

CL974686

CL974686.1

GI:52403885

GSS.

Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

1 (bases 1 to 3048)

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Jiao, X., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,

Wong, G. K. S., Deng, X. W. and Wang, J.

An analysis of transcriptional regulation of the rice genome and

its comparison to Arabidopsis

Unpublished (2004)

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559
 Fax: 86-10-80488676
 Email: chenchen@genomics.org.cn
 Rice genomic sequence.
 Class: exon-trapped.
 Location/Qualifiers
 1. 3048
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="genomic DNA"
 /db_xref="taxon:39946"
 /clone_lib="Oryza sativa Express Library"
 /note="Oryza sativa exon trapped genomic sequences"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,77e-197 Length: 3048
 Score: 1982.00 Matches: 439
 Percent Similarity: 55.29% Conservative: 157
 Best Local Similarity: 40.72% Mismatches: 281
 Query Match: 35.09% Indels: 201
 DB: 9 Gaps: 24
 US-10-627-132-30 (1-1052) x CL974686 (1-3048)
 QY 75 GlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLys----- 90
 Db 15 GGGGAAGAGGGGATCCGGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 74
 QY 91 GlySerProArgValAlaGlyAspAspAspGluGluAspIleAspAspLeuGluHisGlu 110
 Db 75 CGGGTGGCGG--- 131
 QY 111 PheAsnIleAspAspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGln 130
 Db 132 -----GCG 134
 QY 131 IleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArg----- 144
 Db 135 GTACGTGACCTGTCCGGGGAGACATCGACATGAGGCGGAGCTGGCGGCGGAGTACAC 194
 QY 145 -----GlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProIlelle 161
 Db 195 CAACTACAGGTGCAGATCCCGCGAGCGCCGACAA-----CCAGCCGAT 239
 QY 162 ThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHis 181
 Db 240 GCTGAACGCGCGGAGCGG--- 287
 QY 182 GlyGluValSerSerSerLeuHisLysArgIle-----HisProTyrProValSer 198
 Db 288 GTACGTGTCCAACTCCCTCTTACCGGGGCTTCAACAGCGCCACCGCGCCACCTCAT 347
 QY 199 GluProGly----- 201
 Db 348 GGACAAGGTATCGAGTCGTGGTCTCCACCGCGAGATGGCGCGCGCGCGCGCGCGCGG 407
 QY 202 -----SerAlaLysTrpAspGluLysLysGluValSerTrpLysGluArg----- 216
 Db 408 CTGCGCCATCCCGCTGCGAGCGGAGCCGATGCGCAACGAGCGCGCGCGCGCGCGCGCG 467
 QY 217 ---MetAspAspTrpLysSerLysGln-GlyIleLeuGlyGlyGlyAlaAspProGluAs 235
 Db 468 ACCCTGCGAACTTGGCGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 525
 QY 235 pMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSe 255
 Db 526 -----CAGAGCGGTTCAGCGCGCTCAGCGCGCTGAGCGCGCGCGCGCGCGCGCG 560
 QY 255 rIleAlaSerSerLysValAsnProTyrArgMetValIleValValArgLysValVal 275
 Db 561 GATCGCGAGCTGGTGATCAGCCCGTACCGGATCTTCATGTCGATCCGATGTCGTCGT 620
 QY 275 uAlaPheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLe 295

Db 621 CCTTCTTACCTCAGTGGCGCATCGCAACCAACATGGAGGCGCTCTGCTGTGGGG 680
 QY 295 uValSerIlelleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLy 315
 Db 681 GATGTCCATCGTCTGGAGCTCTGTTGGCTTCTCATGGCTGCTCGCATGCTCCCCAA 740
 QY 315 sTrpPheProIleAspArgGluThrTyrIleuAspArgIleuSerLeuArgGlu 335
 Db 741 GGTGAACCGGTCAACCGGAGCACCGCTCCCGTGTCTCAAGGAGAAGTTCGAGACGCC 800
 QY 335 uGlyGluPro-----SerLeuLeuSerAlaValAspLeuPheValSerth 350
 Db 801 GTCCGCGTCCAAACCGCGCGCGCTCCGACCTCCCGGCGCTCGACGCTTCTGCTCCAC 860
 QY 350 rValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAl 370
 Db 861 CGCCGATCGGAGAGAGCGCGTCTCACACGCGGAGCACCATCTCTCCATCTCTCGC 920
 QY 370 aValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLe 390
 Db 921 CGTGACTACCCCGTCCGAGAGCTCGCTGTCTACGTCGACGAGCGCGCGCGCGTGT 980
 QY 390 uThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTyrValProPheCysLy 410
 Db 981 CACCTTCGAGGCGCATGGCGGAGCGCGCGAGCTTCGCCAACGCTGGGTGCTCTTGC 1040
 QY 410 sLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrIle 430
 Db 1041 GAAGCAGCACATCGAGCCCGCCGACCCCGAGCAGCTACTTTCAGCTCAGGCGGACCCG 1100
 QY 430 uLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGl 450
 Db 1101 GAAGGCGAAGCGCGCAACGACTTCGTCAAGGACCGCGCGAGGCTGAAGCGCGAGTTCGA 1160
 QY 450 uGluPheLysValArgIleAsnAlaLeu----- 459
 Db 1161 CGAGTTCAAGGTGGCATCAATGGCTCCCGACTTCATCCGCGCGCGCTCCGACGCTT 1220
 QY 460 -----ValAl 461
 Db 1221 CAACGCGCGGAGACATGAAGATGCTCAAGCACCCTCCGCGAGACGGCGCGCGCGCTC 1280
 QY 461 aLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProGl 481
 Db 1281 GGAGCAGCCCAAGGTGAAGAAGAGCGGACGCTGG---ATGGCGGACGCTCGCATCGCGCG 1337
 QY 481 Y-----AsnAsnThrArgAspHisProGlyMetIleGluVa 493
 Db 1338 GACATGGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCAG 1397
 QY 493 lPheLeuGlyHisSer-----GlyGlyHisAspThrGlu----- 504
 Db 1398 GATGCTCAAGCGCGCGCTCGCGCGCGCGCGCTCTACGGGATGACGACGAGCACCATGAT 1457
 QY 505 -----GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysAr 519
 Db 1458 CGACTTCAGCGAGCTCGACATCCGCTCCCGATGCTCTACATGTCGCGCGGAGAGCG 1517
 QY 519 gProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaIleuIleArgValSerAl 539
 Db 1518 CCCGGCTTACGACCAACAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1577
 QY 539 aValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTrpIleAsnAsnSe 559
 Db 1578 CGTATGTCGAGACGGCCCTTCATGCTCACTTCGACTGCCACCTACATCAACAGCG 1637
 QY 559 rLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCy 579
 Db 1638 GCAGCGGTGGCGGAGCGCATGTCTTCTTCATGAC---CGCGCGCGCGGAGCGCATCGC 1694
 QY 579 sTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnAr 599

```
Db 1695 CTACATCAGTTCCGCGCGCTTCGAGGGGATCGACCCCTCCGACCGCTACGCCAACAA 1754
Qy 599 gAsnThrValPhePheAspIleAenMetLysGlyLeuAspGlyIleGlnGlyProValTy 619
Db 1755 CAACACCGCTCTTCTCGATGCAACATCGCGCGCTCGAGCGCTCCAGGGCCCCATGTA 1814
Qy 619 xValGlyThrGlyCysValPheArgArgGlnAlaLeuTyGlyTyzAsnProProLysG1 639
Db 1815 CGTCGGCAGCGCTGCAATGTCGCGCTTCGCGCTCTAGCGCTTCAGACCTCCGCGCAC 1874
Qy 639 yProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLysLysAr 659
Db 1875 CGCCAGTACACGGGTGGCTC-----TTACCAAGAGAGGT 1913
Qy 659 gLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLy 679
Db 1914 GACCAGCTTCAAGACCGGAGTCCGACACCCAGACGCTCAAGCGGGAGGACTTCGACGC 1973
Qy 679 sGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheVa 699
Db 1974 CGAG---CTCAGTCGCACCTC---GTCCCGCGCGGTTCGGCAACTCGTCGCCGTTCAT 2027
Qy 699 lThrSerThrLeuMetGluGluGlyGlyValProPro---SerSerSerProAlaAlaLe 718
Db 2028 GGCCTCGATCCCGTCGCGAGTTCAGGCGCGGCCCTCGCCGACACCCGCGCGGTGCT 2087
Qy 718 u-----Le 719
Db 2088 GCACGGCAGCGCTCCGCGCGCTCACCGTCCCGCGCGCGCTGGACCCGCGACGGT 2147
Qy 719 uLysGluAlaIleHisValIleSerCysGlyTyzGluAspLysThrAspTrpGlyLeuG1 739
Db 2148 GCGCGAGCGCTGTCGGTGATCTCGTGTGTGTACGAGGACAGGAGTGGGGCGGACCG 2207
Qy 739 uLeuGlyTrpIleTyzGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCy 759
Db 2208 CGTGGGTGGATCTACGGCTCGGTACCGAGGAGCTGGTGACCGGCTACCGGATGCACAA 2267
Qy 759 sArgGlyTrpArgSerValTyzCysMetProLysArgAlaAlaPheLysGlySerAlaPr 779
Db 2268 CCGCGGTGGCGCTCCGTCTACTGCATCACCAGCGCGACGCTTCCTGGGCGACGGCGCC 2327
Qy 779 oIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluI1 799
Db 2328 GATCAACTCACGACAGCTCCACAGTGTCTCCGTGGGCGCGGCTCCGTTCGAT 2387
Qy 799 ePhePheSerArgHisSerProLeuLeuTyzGlyTyzLysAsnGlyAsnLeuLysTrpLe 819
Db 2388 CTTCCTCTCCCGAACACAGCGTTCCTGGCGTCGGCAAG-----CTCATGCTGCT 2438
Qy 819 uGluArgPheAlaTyzIleAsnThrIleTyzProPheThrSerLeuProLeuLeuAl 839
Db 2439 GCAGCGGATCTCCTACCTCACGTCGGCATCTACCCCTTCACCTCCATCTTCCTCCTCGT 2498
Qy 839 aTyzCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSe 859
Db 2499 CTACGCTTTCATCCCGCGGCTCTCCCTCTTCCTCCGGCTTCCTTCATCGCCAGAGCTCGA 2558
Qy 859 rThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLe 879
Db 2559 CATCGCTTCTTGTCTACCTCTCACCATGACCATCACCCTCGTCGCCCTCGGCATCCT 2618
Qy 879 uGluMetArgTrpSerGlyValSerIleGluIuTrpTrpArgAsnGluGlnPheTrpVa 899
Db 2619 CGAG----- 2622
Qy 899 lIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAl 919
Db 2623 -----GGGTGCTGAAGGTGATGCG 2642
Qy 919 aGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGlu---AspAspGl 938
Db 2643 CGGGATAGAGATCTCCTTCACTGTCAGCGGAGGCGGCGGCGGCGGCGGACGACAGGAGCAT 2702
```

Search completed: March 12, 2005, 14:26:04
Job time : 5268 secs

```
Qy 938 uPheAlaGluLeuTyzAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLe 958
Db 2703 CTACGGCGACCTTACATCGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2762
Qy 958 uIleIleAsnValIleGlyValValAlaGlyLysSerAspAlaIleAsnAsnGlyTyzG1 978
Db 2763 GATGGTGAACATCATCGCCATCGCTTCGGCTTCGGCGAGGACCATCTACAGCGACAAACC 2822
Qy 978 nSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTy 998
Db 2823 GCGGTGGGGAGATTTCATCGCGCGCGCTTCTTCAGCTTCTGGGTGCTCGCCACCTCAA 2882
Qy 998 rProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTr 1018
Db 2883 CCCCTTCGCCAAGGGCTCATGGGGGGCGCGCGGCAAGACCCACCATCGTCTTCGTCGT 2942
Qy 1018 pSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPro 1035
Db 2943 GTCGGGCTCTCTCAATCACCGTCTCCCTGCTCTGGGTGCGCATCAGCCCG 2994
```

QY 1861 CGACGTGCACGACCGATAGCTAAGCAGGAACACCGCTCTTCTCGACATCAACATGAAGG 1920
Db 1861 CGACGTGCACGACCGATAGCTAAGCAGGAACACCGCTCTTCTCGACATCAACATGAAGG 1920
QY 1921 GCTGACGGCATCCCAAGGCGCGGTGTAGCTCGGGACAGGGTGCCTGTTCCGGCGCCAGGC 1980
Db 1921 GCTGACGGCATCCCAAGGCGCGGTGTAGCTCGGGACAGGGTGCCTGTTCCGGCGCCAGGC 1980
QY 1981 GCTCTACGGCTAACAACCTTCCCAAGGACCCCAAGAGGCCCAAGATGGTGAACCTGCGACTG 2040
Db 1981 GCTCTACGGCTAACAACCTTCCCAAGGACCCCAAGAGGCCCAAGATGGTGAACCTGCGACTG 2040
QY 2041 CTGCCCCGTCTTCCGCGCGCAAGCAAGCGGAACACGCCAAGGACGGCTGCCGGAGGCAC 2100
Db 2041 CTGCCCCGTCTTCCGCGCGCAAGCAAGCGGAACACGCCAAGGACGGCTGCCGGAGGCAC 2100
QY 2101 CGCTGATATGGAGTAGATAGCGACAAGGAGATGCTCATGTCCCAATCACTTCGAGAA 2160
Db 2101 CGCTGATATGGAGTAGATAGCGACAAGGAGATGCTCATGTCCCAATCACTTCGAGAA 2160
QY 2161 GCGGTTCCGGCATCCCGCGCGCTTCTGTCAGTCGACGTCGATGAGAGGAGCGCGCTCCC 2220
Db 2161 GCGGTTCCGGCATCCCGCGCGCTTCTGTCAGTCGACGTCGATGAGAGGAGCGCGCTCCC 2220
QY 2221 TCCTTCGTCGAGCCCGCGCGCTCTCAAGAGGCCATCCATGTATCATCAGTCCGGCTA 2280
Db 2221 TCCTTCGTCGAGCCCGCGCGCTCTCAAGAGGCCATCCATGTATCATCAGTCCGGCTA 2280
QY 2281 CGAGCAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTGCGATCACGGAGGA 2340
Db 2281 CGAGCAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTGCGATCACGGAGGA 2340
QY 2341 CATCTGACGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTTGATCTGATGCCGAA 2400
Db 2341 CATCTGACGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTTGATCTGATGCCGAA 2400
QY 2401 GCGGCGCGGTTCAAGGGTGGCGCGCATCAATCTATCGGACCGTCTCAACCGAGTCT 2460
Db 2401 GCGGCGCGGTTCAAGGGTGGCGCGCATCAATCTATCGGACCGTCTCAACCGAGTCT 2460
QY 2461 CCGGTGGCGCTGGGGTCCGTGAGATCTTCTCAGCGCGGACACGCCCTCTGCTGACGG 2520
Db 2461 CCGGTGGCGCTGGGGTCCGTGAGATCTTCTCAGCGCGGACACGCCCTCTGCTGACGG 2520
QY 2521 CTACAAGAACGGCAACCTCAAGTGGCTGAGCGCTTCGCCCTACATCAACACCAATCTA 2580
Db 2521 CTACAAGAACGGCAACCTCAAGTGGCTGAGCGCTTCGCCCTACATCAACACCAATCTA 2580
QY 2581 CCCCTTCACTCGCTCCCGCTGCTGCTACTGCAACCTCCCGCGCTGCTGCTCTCAC 2640
Db 2581 CCCCTTCACTCGCTCCCGCTGCTGCTACTGCAACCTCCCGCGCTGCTGCTCTCAC 2640
QY 2641 CGCAAGTTCATCATCGCGTTCGATGACGCTTCGCCAGCTCTTCTTCATCGCCCTCT 2700
Db 2641 CGCAAGTTCATCATCGCGTTCGATGACGCTTCGCCAGCTCTTCTTCATCGCCCTCT 2700
QY 2701 CATGTCATCTTCGCGACGGGCATCTCTGAGATGCGGTGGAGCGGGGTGAGCATCGAGGA 2760
Db 2701 CATGTCATCTTCGCGACGGGCATCTCTGAGATGCGGTGGAGCGGGGTGAGCATCGAGGA 2760
QY 2761 GTGTTGAGGACGACGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Db 2761 GTGTTGAGGACGACGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
QY 2821 CGTGACGCGCTCTCAAGTCTCTCGCGGATCGACACCACTTCACCGTCACTCCAA 2880
Db 2821 CGTGACGCGCTCTCAAGTCTCTCGCGGATCGACACCACTTCACCGTCACTCCAA 2880
QY 2881 GGCCACCGGACGAGGACGAGTTCGCGAGCTCTACGCTTCAAGTGGACACCGCT 2940
Db 2881 GGCCACCGGACGAGGACGAGTTCGCGAGCTCTACGCTTCAAGTGGACACCGCT 2940
QY 2941 CCTCATCCCGCCACCGCTGCTCATATTAACGTATCGGGCTGCTGGCGGATCTC 3000

Db 2941 CCTCATCCCGCCACCGCTGCTCATATTAACGTATCGGGCTGCTGGCGGATCTC 3000
QY 3001 CGAGCCATCAACAAAGGCTACAGTCTCTGGGGGCCCTCTTTCGCGAAGCTCTTCTTCGC 3060
Db 3001 CGAGCCATCAACAAAGGCTACAGTCTCTGGGGGCCCTCTTTCGCGAAGCTCTTCTTCGC 3060
QY 3061 CTTCTGGGTCTAGTCTCACTCTACCGCTTCTCAAGGGGCTCATGGGGCGCGAAGACAG 3120
Db 3061 CTTCTGGGTCTAGTCTCACTCTACCGCTTCTCAAGGGGCTCATGGGGCGCGAAGACAG 3120
QY 3121 GACGCCCAAGCTGCTGCTCATCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
Db 3121 GACGCCCAAGCTGCTGCTCATCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
QY 3181 GGTGAGGATCAGCCCTTTCATCGTCAGGACCAAGGGCCCGGACGTCAGGACGATGCGCAT 3240
Db 3181 GGTGAGGATCAGCCCTTTCATCGTCAGGACCAAGGGCCCGGACGTCAGGACGATGCGCAT 3240
QY 3241 CAATTGCTGAGCTGTTTAAAGGTTCAAAATTCCTGAGCTTGTGCATAGGAGAGAAAA 3300
Db 3241 CAATTGCTGAGCTGTTTAAAGGTTCAAAATTCCTGAGCTTGTGCATAGGAGAGAAAA 3300
QY 3301 ACATTTAGAAATTTTGAAGGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
Db 3301 ACATTTAGAAATTTTGAAGGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
QY 3361 GACGAGGAATTTGAAGGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
Db 3361 GACGAGGAATTTGAAGGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
QY 3421 AAAAAA 3443
Db 3421 AAAAAA 3443

RESULT 2

AK121170
LOCUS
DEFINITION
ORyza sativa (japonica cultivar-group) cDNA clone: J023081B08, full insert sequence.
ACCESSION
AK121170
VERSION
AK121170.1 GI:37990793
KEYWORDS
FLI_CDNA; CAP trapper.
SOURCE
ORyza sativa (japonica cultivar-group)
ORyza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, B., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ohtsuki, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tanoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narioka, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kusumegi, T., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oosato, N., Oca, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
TITLE
japonica rice
JOURNAL
Science 301 (5631), 376-379 (2003)
MEDLINE
22752273
PUBMED
12869764
REFERENCE
2

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice

Unpublished
3 (bases 1 to 3631)

JOURNAL

Kikuchi, S.

Direct Submission

REFERENCE

Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp).

TS:81-29-838-7007, Fax:81-29-838-7007

This clone is one of the 32K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NTAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M. and Nakahama, Y.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Iehibiki, J., Kawamata, M.

Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., MacCauyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ooka, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

1. 3631
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023081B08"

ORIGIN

Query Match

Beat Local Similarity 78.2%; Score 2691.2; DB 8; Length 3631;

Matches 2953; Conservative 0; Mismatches 323; Indels 18; Gaps 5;

Qy	67	CCTAGCACGCGCGCCTCCGCGCGCGATGGAGGCGCAGCGCCGGGCTGGTGGCCGGCTCGCA	126
Db	71	CGTTGCGCGCGCTGTGTTCGCGCGCCATGGAGGCGAGCGCGCGGCTGGTGGCCGGCTCGCA	130
Qy	127	CRACCCGACGAGCTGGTCTGTATCCGGGCGCAGAGGACCCCAAGCGCTGGCGGCGCT	186
Db	131	CAACCCGACGAGCTGGTCTGTATCCGGGCGCAGAGGACCCCAAGCGCTGGCGGCGCT	190
Qy	187	GAGCGGCGAGGTGTGGAGATATGCGCGCGCAGAGGTCGGGCTCACGGTGACGCGCGACCT	246
Db	191	GAGCGGCGAGGTGTGGAGATATGCGCGCGCAGAGGTCGGGCTCACGGTGACGCGCGACCT	250
Qy	247	CTTTCGTCGCTGCAACGAGTGGCGCTTCCCGTGTGCGCGCCCTGTCTACGATACGAGCG	306
Db	251	CTTTCGTCGCTGCAACGAGTGGCGCTTCCCGTGTGCGCGCCCTGTCTACGATACGAGCG	310
Qy	307	CCGGGAGGCGACGACGACCTGCCCGCAGTGCAGAGCGCTACAGCGCTCCTCAGGGGAG	366
Db	311	CCGGGAGGCGACGACGACCTGCCCGCAGTGCAGAGCGCTCCTCAGGGGAG	370
Qy	367	CCCGAGGGTTGCGCGGCGACGATGACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAA	426
Db	371	CCCGAGGGTTGCGCGGCGACGATGACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAA	430
Qy	427	CATCGACGACGAGAAATCAGCAGAGGAGGAGCTGGAGGGCAA-----CATGACGAAACAGCCA	480
Db	431	CATCGACGACGAGAAATCAGCAGAGGAGGAGCTGGAGGGCAA-----CATGACGAAACAGCCA	490
Qy	481	GATCAGCGAGGCGATGCTGACGCGGAGGATGAGTACGGGAGGGGCCCCGACGACGCGGA	540
Db	491	CATCAGCGAGGCGATGCTGACGCGGAGGATGAGTACGGGAGGGGCCCCGACGACGCGGA	550
Qy	541	CGCGCAACAAACACCCCGCAGATCCCGCCATCATCACCGGCTCCCGCTCCCGGCTGAG	600
Db	551	CGCGCAACAGACCCCGC---TCCCGCGATCATCACCGGCGCTCGCTCCCGGCTGAG	607
Qy	601	CGGTGATTTTCCGATTTACCAACCGGATGATGCCACGGGAGGTTCGTTTCCCTGCAAAA	660
Db	608	CGGGGAGTTTCCGATTTACCAACCGGATGATGCCACGGGAGGTTCGTTTCCCTGCAAAA	667
Qy	661	GGCCATCCATCCCTGTCCTGTCGTCGACGAGGAGTGCCCAAGTGGGACGAGGAAGA	720
Db	668	GGCCATCCATCCCTGTCGTCGACGAGGAGTGCCCAAGTGGGACGAGGAAGA	727
Qy	721	AGTGAGCTGGAAGGAGAGATGGAACGACTGGAAGTCCAAAGCAGGGCATCTCCGCGCGG	780
Db	728	GGTGAGCTGGAAGGAGAGATGGAACGACTGGAAGTCCAAAGCAGGGCATCTCCGCGCGG	787
Qy	781	CG---CCGATCCCGAAGACATGGAACGCGCGAGCTGGGCACTGAAACGAGCGGAGGAGCC	837
Db	788	CGCCCCCGATCCCGACGACTACGACGCGCGACGCTCCCACTGAACGACGAGGCGAGGAGCC	847
Qy	838	GCTGTGAGGAAGGTGTCCGATCCCGTCGACGAGGTGAACCCGTCACCGATGCGGATCGT	897
Db	848	GCTGTGAGGAAGGTGTCCGATCCCGTCGACGAGGTGAACCCGTCACCGATGCGGATCGT	907
Qy	898	GGTGGCTCTCGTTGTGCTCGCCCTTTCTCTCCGGTACCGGTATCTGTCACCCCGCGGGA	957
Db	908	CCTCCGCTCTCGTTGTGCTCGCCCTTTCTCTCCGGTACCGGTATCTGTCACCCCGCGGGA	967
Qy	958	CGCCATCGGGCTGTGGCTGTCTTCATCATCTGCGAGATCTGGTTGCGCATCTCTCGAT	1017
Db	968	CGCCATCGGGCTGTGGCTGTCTTCATCATCTGCGAGATCTGGTTGCGCATCTCTCGAT	1027
Qy	1018	CCTCGACGAGTCCCAAGTGGTTTCCCATCGACGCGGAGACGCTACCTCGACCGCTCTC	1077
Db	1028	CCTCGACGAGTCCCAAGTGGTACCCGATCGAGGAGACCTTACCTCGACCGCTCTC	1087
Qy	1078	CCTCAGGTACGAGGAGGAGGAGCGCTCGCTGCTGTTCGGCGGTGGACCTGTTGCTGAG	1137
Db	1088	CCTCCGCTACGAGCGGAGGAGGAGCGCTCGCTGCTGTTCGGCGGTGGACCTGTTGCTGAG	1147

D	b	7123	TACCGGATGGTGATCATCTCTCGTGTGTCGCGCTTCTTCCTCCCGTACCGCATC	7182	Qy	1850	TTCCGGCGCATCGACGTGCACGACCGATACGCTAACAGGAACACCGCTCTTCTTCGAC---	1906	
Q	y	941	CTGCACCCCGTCCCGGACCGCATCGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGG	1000	D	b	8263	TTCCGACGGCATCGACGTTCACGACCGGATACGCCAACCGCAACACCGCTCTTCTTCGACGTG	8322
D	b	7183	CTCCACCGGTGCCGAGCGCATCCCGTGTGGTTCACCTCCATCATCTGCGAGATCTGG	7242	Qy	1907	-----	-----	1906
Q	y	1001	TTCCGCATCTCTCGATCTCTCGACCAAGTTCCTCCCAAGTGGTTCCTCCATCGACCGGAGACG	1060	D	b	8323	AGCTCTCTCCACACAACTAGATGAATATATACAATCTTGAATAATCCAGAGCAATCA	8382
D	b	7243	TTCCCGGTGTCTCGATCTCTCGACCAAGTTCCTCCCAAGTGGTACCCCGATCGACAGGAGACC	7302	Qy	1907	-----	-----	-A 1907
Q	y	1061	TACCTCGACCGCTCTCCCTCAGTACGAGAGGAAAGGAGCGCTGCTGCTGTCTCGCG	1120	D	b	8383	CGATCGATGAGCAATGTGACTGAATTTTGTGTGTGCGTCTTGTGTGAGATCATCAGA	8442
D	b	7303	TACCTCGACCGCTCTCCCTCAGTACGAGCGAGAGGAGCGCTGCTGCTGTCTCGCG	7362	Qy	1908	TCACATGAAGGGCTCGACCGCATCCAAAGCCCGGTGTACGTCTGGGACAGGGTGTCTGT	1967	
Q	y	1121	GTGACCTGTTCGTGAGCACGGTGGACCCCGTCAAGGAGCGCGCTGTGTGACCGCAAC	1180	D	b	8443	TCACATGAAGGGCTGTGATGGATCCAGGGCCGGTGTACGTCTGGGACAGGGTGTCTGT	8502
D	b	7363	GTGACCTGTTCGTGAGCACGGTGGATCCCGTCAAGGAGCGCGCTGTGTGACCGCAAC	7422	Qy	1968	TTCCGGCGCAGGCGCTCTACGGCTACAACTCTCCAAAGGACCCAGAGGCCCAAGATGG	2027	
Q	y	1181	ACCGTGTCTCCATCTCTCGCGGTAGACTACCCCGTGGACAAAGTGTCTCTGCTACGTCTCC	1240	D	b	8503	TCAGGGCGGAGGCGCTGTACGGATACAAACCCCAAGGGACCCAGAGGCCCAAGATGG	8562
D	b	7423	ACCGTGTCTCCATCTCTCGCGGTGCACTACCCCGTGGACAAAGTGTCTCTGCTACGTCTCC	7482	Qy	2028	TGACCTGCGACTCTGCGCGTGTCTCGCGCGCAAGAGCGGAAACACGCCAAAGACGGGC	2087	
Q	y	1241	GACGACGGCGGTGATGCTGACCTTCGAGTCTGCTGTGGAGACGGCGGAGTTCGCGGC	1300	D	b	8563	TGACCTGCGACTCTGCGCTTGTCTTCCGGAGGAGAGCGGAAAGCAGCGGACGCGCC	8622
D	b	7483	GACGACGGCGGTGATGCTGACCTTCGAGTCTGCTGTGGAGACGGCGGAGTTCGCGGC	7542	Qy	2088	TGCCGGAGGGCACCGC-----	-----	2103
Q	y	1301	AAGTGGGTGCCCTTCTGCAAGAGTTCGGCATCGAGCCCGCGCCCGGAGTTCATCTTC	1360	D	b	8623	TCCCGGAGGCGCTCGCGCGACGGCGGTGAGCTCCCAAAATTCAGAGCTAAGAAGAAAT	8682
D	b	7543	AAGTGGGTGCCCTTCTGCAAGAGTTCAGCATCGAGCCCGCGCCCGGAGTTCATCTTC	7602	Qy	2104	-----	-----	TGATATGGGA 2113
Q	y	1361	TCGCTCAAGGTGCTACCTCAAGGACAAAGTTCAGGCTCAGGCGGAGCGCGC	1420	D	b	8683	ATGGTGTGAAGATTTTCAGCTGATGAATAATTAAGTTTGTGTGTGGATCAGGG	8742
D	b	7603	TCCAGAGAGTTCGACTACCTCAAGGACAAAGTTCATCCCACTTCGTCAGGAGCGCGC	7662	Qy	2114	GTAGATAGACAGAGAGATGTCTCATGTCCCATCATGAATTTCCAGAGCGGCTCGGCGAG	2173	
Q	y	1421	GCCATGA-----	1427	D	b	8743	ATGGACAGCGACAAAGAGATGTCTCATGTCCGATGAATTCGAGAGCGGTTTCGGGCG	8802
D	b	7663	GCCATGAAGTAAATATCGATCGCATCGTCAATGCTGATTTCTGTAGAGCTTTGACGAAGA	7722	Qy	2174	TTCCCGCGCTTCTGCTACGCTCGACGCTGATGGAGGAAGCGGCGCTCCCTCTCTGCTCGAGC	2233	
Q	y	1428	-----	1451	D	b	8803	TTCCCGCGCTTCTGCTGACGCTCGACGCTGATGGAGGAAGCGGCGCTCCCGCGCTCGTCCAGC	8862
D	b	7723	GCTAACTGTTGTGGGGCATTTGGCATTTGATCGAGCAGAGAGATACGAGGATTTCAAGG	7782	Qy	2234	CCCGCGCGCTCTCTAAAGGAGCGCATCATGTCTCATGCTGCGGCTACGAGGACCAAGACC	2293	
Q	y	1452	TCCGGATCAACGCGCTGTGTGCCAAGGCTCATGAGTTCGCGGACAGAGGGTGTGATCATGA	1511	D	b	8863	CCCGCGCGCTCTCTAAAGGAGCGCATCATGTCTCATGCTGCGGCTACGAGGACCAAGACC	8922
D	b	7783	TGAGGATCAACGCGCTGTGTGCGAAGCGCAGAAAGGTGCCGCGGAAAGGGTGGATCATGA	7842	Qy	2294	GACTGGGGGCTGG-----	-----	2306
Q	y	1512	AGGACGCAACGCGTGGCGCGGAAACACACCCCGGACCAACCCCGGCGATGATCCAGGTGT	1571	D	b	8923	GACTGGGCTCTCGAGGTAAACCGGTTATCAGACACTAAGCCATATATTACATTTGAGTGAG	8982
D	b	7843	AGGACGCAACGCGTGGCGCGGAAACACACCCCGGACCAACCCCGGCGATGATCCAGGTGT	7902	Qy	2307	-----	-----	AGCTGGGGTGGATCTACGGGTCCGATC 2332
Q	y	1572	TCCTGGGCGCACAGCGGGCGGACACACCGAGGGCAACGAGCTGCCCGCGCTCGTGTACG	1631	D	b	8983	TTTGTGTGTGAACGCCATGGAATGTGTGATGCAGCTGGGGTGGATCTACGGGTCCGATC	9042
D	b	7903	TCCTGGGCGCACAGCGGGCGGACACCGAGGGCAACGAGCTGCCCGCGCTCGTGTACG	7962	Qy	2333	ACGAGGACATCTGACGGGTTCAAGATGCATCTGCGGGGTGGCGCTCCGTGTACTGC	2392	
Q	y	1632	TCTCCCGTGAAGCGCGCGGATTCAGACACCAAGAGAGCGCGCGCCATGAAGCGTCC	1691	D	b	9043	ACGAGGACATCTTAAAGGGGTTCAAGATGCATCTGCGGGGTGGAGGTTCGGTGTACTGC	9102
D	b	7963	TCTCCCGTGAAGCGCGCGGCTTCAGACACCAAGAGAGCGCGCGCCATGAAGCGCC	8022	Qy	2393	ATCCGAGCGCGCGGTTCAAGGGGTGGCGCGCATCAATCTATCGGACCGCTCTCAAC	2452	
Q	y	1692	T-----	1692	D	b	9103	ATCCGAGAGGGCGCGGTTCAAGGGGTGAGCGCGCATCAACCTATCTGACCGCTCTCAAC	9162
D	b	8023	TCGTACGCCACCGCACCATCTCTTCACTTTCAAAAACAACTCTGCTCGATTTGAT	8082	Qy	2453	CAGGTGCTCCGGTGGGCGCTGGGGTCCGTGAGATCTTCTTACGCGGGACAGCGCCCTG	2512	
Q	y	1693	-----	1729	D	b	9163	CAGGTGCTCCGGTGGGCGCTCGGCTCCGTGAGATCTTCTTACGCGGGACAGCGCCGCTC	9222
D	b	8083	CGAAATTTCTGGTTTGGCAGATTCGTGTGTGCGCGGTGCTGACGAGCGCGCGTTC	8142	Qy	2513	CTGTACGGCTACAAGAACCGCAACCTCAAGTGTGAGGCGCTTCGGCTACATCAACACC	2572	
Q	y	1730	ATGCTCAACTTTGATGTGATCATCATCAACCAACCAAGAGGCGCATCCCGGAGGCCATG	1789	D	b	9223	CTCTACGGCTACAAGAACCGCAACCTCAAGTGTGAGGCGCTTCTCTACATCAACACC	9282
D	b	8143	ATGCTCAACTTTGATGTGATCATCATCAACCAACCAAGAGGCGCATCCCGGAGGCCATG	8202	Qy	2573	ACCATCTACCCCTTACCTCGCTCCGCTCGCTACTGACCCCTCCCGCGGCTCTGC	2632	
Q	y	1790	TGCTTCTCATGACCCCTCAGTTCGGCGGAGGCTCTGCTACGTTCAAGTTCCCGCAGAGG	1849	D	b	9283	ACCATCTACCCCTTACCTTCTCTCCCTCTCTGCGCTACTGACCCCTTACCGCGGCTCTGC	9342
D	b	8203	TGCTTCTCATGACCCCTCAGTTCGGCGGAGGTTTGTACGTCGAGTTTCCCGCAGAGG	8262					

Qy	382	GGACGATGACGAGGAGCAATCGACGACCTGGAGCAGCATGTTCAAATCATCGACGACGAGA	441
Db	121133	GGACGAGGAGGAGGAGCAATTGACGACCTGGAGCACGAGTTCAAATCATCGACGACGAGA	121192
Qy	442	TCAGCAGAGCAGCTGGAGGCA-----CATGCAGAACGCCAGATCAACGAGGCGAT	495
Db	121193	GCGAAGCAGCTCGACGAGATCAGATGGCATGCAAGATGCAAGAACGCCATCATCCGAGGCGAT	121252
Qy	496	GCTGCACGGCAGGATGAGCTTACGGGAGGGGCCCGACGACGCGGACGGCAACAACACCCC	555
Db	121253	GCTGCACGGCAAGATGAGCTTACGGGAGGGGCCCGACGACGCGGACGGCAACGACACCCC	121312
Qy	556	GCGATCCCCGCCCATCATCACCGGCTCCGCTCCGTFGCC-----	594
Db	121313	GC---TCCCGCCGATCATCACCGGCGCTCGCTCCGTFCCCGGTACACACACACACTCAC	121369
Qy	595	-----	594
Db	121370	CCCACTCACACAAAATCTCTCCCTCTTCTCACTCAACATGTTCAAGAAATGTTCTTTTG	121429
Qy	595	-----GGTAGCGGTGAGTTTCCGATTTACCACGGGTATGGCCACGGCGAG	640
Db	121430	TAAAAAAAATTTCAGGTGAGCGGGAGTTCCCGATATCGAACAGCCATGGCCATGGCGAG	121489
Qy	641	GTCTCGTCTTCTGCACGAAGCGCATCCATCCGTACCTCTGTCTGAG-----	688
Db	121490	TTCTCTCTTCTCCCTGCACGAAGCGCATCCACCCCTACCCGGTGTCTGAGCCAGGTAGTATA	121549
Qy	689	-----	688
Db	121550	ACGAATTCACACACTAAATTAATCAATGTTCTTGATGAAACACACATTTGATCATCTCAATC	121609
Qy	689	-----CCAGGGAGTCCCAAGTGGGACGAGAGAAAGAGTAGGCTG	729
Db	121610	ATCTACCGATGAATCTGACCGAGGAGTGCAAGTGGGACGAGAGAAAGGTGAGCTG	121669
Qy	730	GAAGGAGAGATGGACGACTGGAAGTCCAAGCAGGECATCTCGCGCGCGCG---CCGA	786
Db	121670	GAAGGAGAGATGGACGACTGGAAGTCCAAGCAGGECATCTCGCGCGCGCGCGCCCGCA	121729
Qy	787	TCCGGAAGCATGGACGCGAGTGGCACT-----	816
Db	121730	TCCGGAAGCATGGACGCGAGTGGCACTTGACGCAATCTTCAGCTGAGCAGCTTACAC	121789
Qy	817	-----GAAC	820
Db	121790	TGATCATGCATGACAACTAGTATATTGATCATGCCGTGAATCTCTGTGGCTTGACGAAAC	121849
Qy	821	GACGAGCGAGGAGCGCGCTGTGAGGAAGGTGTGATCGCGTGCAGCAAGGTGAACCG	880
Db	121850	GACGAGCGAGGAGCGCGCTGTGAGGAAGGTGTGATCGCGTGCAGCAAGGTGAACCG	121909
Qy	881	TACCGATGGTATCGTGGTGGTCTGCTGTGCTCGCCTCTTCTCGGTACCGTATC	940
Db	121910	TACCGATGGTATCGTGGTGGTCTGCTGTGCTCGCCTCTTCTCGGTACCGTATC	121969
Qy	941	CTGCACCCCTCCGAGCGCATCGGCGTGGCTGCTCTCCATCATCTGCGAGATCTGG	1000
Db	121970	CTGCACCCGCTGCCGAGCGCATCCGCTGTGGCTCATCTTCATCATCTGCGAGATCTGG	122029
Qy	1001	TTCCCATCTCTCTGGATCTCTGACCACTGTTCCCAAGTGGTTCCTCCATCGACCGGAGACG	1060
Db	122030	TTCCGCTGTCTGGATCTCTGACCACTGTTCCCAAGTGGTTCCTCCATCGACCGGAGAC	122089
Qy	1061	TACTCGACCGCTCTCTCTCAGTACGAGAGGAAGGGAGCGCTCGCTGCTGTCGGCG	1120
Db	122090	TACTCGACCGCTCTCTCTCAGTACGAGAGGAAGGGAGCGCTCGCTGCTGTCGGCG	122149
Qy	1121	GTGACCTGTTCTGTGAGCAGGTGGACCGCTCAAGGACCGCGCTGTGACCGCCCAAC	1180
Db	122150	GTGACCTGTTCTGTGAGCAGGTGGATCCGCTCAAGGACCGCGCTGTGACCGCCCAAC	122209
Qy	1181	ACGGTGTCTCTCACTCTCGCGGTAGACTACCCCGTGGACAAGGTTCTCTCTGTAAGTCTCC	1240

Db 123290 TCAGCGGCGGCGCTGTACGGATACAA CCCACCCAAAGGACCCAAAGAGGCCCAAGATGG 123349
Qy 2028 TGACCTGCGACTGCTGCGCGCTGCTCGCGCGCAAGAGCGGAAACACGCGCAAGCGGGC 2087
Db 123350 TGACCTGCGACTGCTGCGCGCTGCTCGCGGAGGAAGCGGAGACGCGCAAGCGGCGC 123409
Qy 2088 TGC CGGAGGCGCACCGC----- 2103
Db 123410 TCCCGAGGCGCTGCGCGCGCGCGGCTGCCAAATTCAGAGCTAAGAAGAAATTT 123469
Qy 2104 -----TGATATGGGA 2113
Db 123470 ATGGTGTGAGAAGATTTTCAGCTGATGGAATATAGTTTGTGTGTGATCAGGG 123529
Qy 2114 GTATAGTAGGACAAAGAGATGCTCATGTCCCAATGAACCTTCAGAGAGCGGTTTCGGGCAG 2173
Db 123530 ATGGACAGCGACAAAGAGATGCTCATGTGCGAGATGAACCTTCAGAGAGCGGTTTCGGGCAG 123589
Qy 2174 TCCGCGCGGCTTCTGACGTGACGCTGATGGAGGAAGGCGCGCTCCCTCTCTGCTCGAGC 2233
Db 123590 TCCGCGCGGCTTCTGACGTGACGCTGATGGAGGAAGGCGCGCTCCCGCGCTCGTCCAGC 123649
Qy 2234 CCGCGCGCGCTCCTCAAGGAGGCGCATTCATGTCTCATGAGCTGCGGCTACGAGGACAAGACC 2293
Db 123650 CCGCGCGCGCTCCTCAAGGAGGCGCATTCATGTCTCATGAGCTGCGGCTACGAGGACAAGACC 123709
Qy 2294 GACTGGGGGCTGG----- 2306
Db 123710 GACTGGGGTCTCGAGGTAAACACCGTTATCAGACACTAAGCCATATATACCATTTGAGTGAG 123769
Qy 2307 -----ACGTGGGGTGGATCTACGGGTGCGATC 2332
Db 123770 TTTGTGTGTGAACGCGCATGGATGTGTGATGCAGCTGGGGTGGATCTACGGGTGCGATC 123829
Qy 2333 ACGGAGGACATCTGACGCGGTTCAAGATGSCACTGCGCGGGTGGCGCTCCGNGTACTGC 2392
Db 123830 ACGGAGGACATCTTAACGGGTTCAAGATGSCACTGCGCGGGTGGAGGTCGGGTGACTGC 123889
Qy 2393 ATCCGAAGCGGCGGCTTTCAAGGGGTGCGGCGCGATCAATCTATCGGACCGCTCTCAAC 2452
Db 123890 ATGCCGAAGAGGCGCGGTTCAAGGGGTGAGCGCGCATCAACCTATCTGACCTCTCAAC 123949
Qy 2453 CAGGTGCTCGGTGGGCGCTGGGGTCTGCGAGATCTTTTCAGCGCGGACAGCCCGCTG 2512
Db 123950 CAGGTGCTCGGTGGGCGCTCGGCTCGGTGCGAGATCTTTTCAGCGCGGACAGCCCGCTC 124009
Qy 2513 CTGTACGGCTTACGAAGCGCAACCTCAAGTGTGCGAGCGCTTCGCTTACATCAACACC 2572
Db 124010 CTGTACGGCTTACGAAGCGCAACCTCAAGTGTGCTGAGCGCTTCTCTACATCAACACC 124069
Qy 2573 ACCATCTACCCCTTACCTCGCTCCGCGTCTCGCCTACTGCAACCCCTCCCGCGCTGTC 2632
Db 124070 ACCATCTACCCCTTCACTTCTCTCCCTCTCGCTACTGCAACCTACCGCGCTGTC 124129
Qy 2633 CTCTCACCGGCAAGTTTCATATGCC----- 2658
Db 124130 CTCTCACCGCAAGTTTCATATGCC----- 124189
Qy 2659 ----- 2658
Db 124190 CTCTCTTCTTCATGATTTTCCCGCGCAAACTGAAGTTTCAAAATGTTCTGAACCTTTT 124249
Qy 2659 -----GTGATTAGCAGTTTCGCGAGCGCTCTTCTTCATCGCCCTCTTTCATGTCCAT 2709
Db 124250 CTTGGTGATGTCAGATTAGCAGGTTTGGAGTTTGTCTTCTTCATCGCGCTCTTCATCTCCAT 124309
Qy 2710 CTTTCGCGACGGGACATCTTGAGATGCGGTGGAGCGGGGTGAGCATCGAGAGTGTGGAG 2769
Db 124310 CTTTCGCGACGGGACATCTTGAGATGAGGTGGAGCGGGGTGAGCATCGAGAGTGTGGAG 124369
Qy 2770 GAACGAGCAGTTCTTGGGTCTATCGGCGCGGTGTCGCGCATCTCTTCGCGTCTGTCAGGG 2829
Db 124370 GAACGAGCAGTTCTTGGGTCTATCGGCGCGGTGTCGCGCGCACCTGTTCGCGGTGTGAGGG 124429

Qy 2830 CTTGCTCAAGGTCTCTGCGCGGATCGACACCAACTTTCACCGTCACCTCCAAGGCCACCGG 2889
Db 124430 CTTGCTCAAGGTGCTGGCGGGATCGACACCAACTTTCACCGTCACCTCCAAGGCCACCGG 124489
Qy 2890 CGACGAGGACGAGAGTTCGCGAGCTCTACGGCTTCAAGTGGACACGACTCTCTATCCC 2949
Db 124490 AGACGAGGACGAGAGTTCGCGAGCTCTACGGCTTCAAGTGGACACGACTCTCTATCCC 124549
Qy 2950 GCCACACGCTGCTCATCATTAACGTCATCGCGCTCGTGGCGGATCTCCGACGCCAT 3009
Db 124550 GCCACACGCTGCTCATCTCAACATCATCGCGTCTGCGCGGCTCTCCGACGCCAT 124609
Qy 3010 CAAACACGGGTACCAAGTCTCTGGGGCCCTCTTTCGGCAAGCTCTTCTTCGCTTCTGGGT 3069
Db 124610 CAAACACGGCTCCGAGGCGTGGGGCCGCTCTTCGGGAAGCTCTTCTTCGCTTCTGGGT 124669
Qy 3070 CATCGTCCACCTCTACCCGCTTCTCAAGGGGCTCATGGGGCGCCAGAACGAGAGCGCCAC 3129
Db 124670 CATCGTCCACCTCTACCCGCTTCTCAAGGGGCTCATGGGGAGGCGAGAACGAGAGCGCCAC 124729
Qy 3130 CGTTGTTGTTCATCTGGTCCATTTCTGCTGGCTCCATCTTCTCCTGCTCTGGGTCAAGAT 3189
Db 124730 CATTTGTTGTTCATCTGGTCTGCTGCTCGCTCCATCTTTTCTTCTGCTCTGGGTCAAGAT 124789
Qy 3190 CGACCCCTTTCATCGTCAGGACCAAGGGCCCGAAGTTCAGGCAAGTGTGGCATCAATTTGCTG 3249
Db 124790 TGATCCCTTTCACCATCAAGGCGAGGGGCGCTGACGTCAGGCGAGTGGCGCATCACTGCTG 124849
Qy 3250 AGCTGTTTATTAAGGTTCAAAATTCGGAGCTTGTGCATAGGAGAGAGAGAAACAATTTAG 3309
Db 124850 AGAGAGGGCGCTCAG---CGATTTCTGAAGATTTGTGCATAGGGGGCGAGACAGATCG 124906
Qy 3310 AAATTTGTAAAGTTGTTGCTGCTGCTGTAATGTTATGTGATC 3348
Db 124907 ATTTGTAAAAGTTTGTATTTGCTGCTGTTGAGTTCGTGC 124945

RESULT 5
AY095297
LOCUS AY095297 3277 bp mRNA linear PLN 07-OCT-2002
DEFINITION Populus tremuloides cellulose synthase (CesA2) mRNA, complete cds.
ACCESSION AY095297
VERSION AY095297.1 GI:23534478
KEYWORDS
SOURCE
ORGANISM
Populus tremuloides (quaking aspen)
Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 3277)
AUTHORS Samuga,A. and Joshi,C.P.
TITLE A new cellulose synthase gene (PtrCesA2) from aspen xylem is
orthologous to Arabidopsis AtCesA7 (irx3) gene associated with
secondary cell wall synthesis
JOURNAL Gene 296 (1-2), 37-44 (2002)
REFERENCE 2 (bases 1 to 3277)
AUTHORS Samuga,A. and Joshi,C.P.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2002) Forestry, Michigan Tech University, 1400
Townsend Drive, Houghton, MI 49931, USA

FEATURES
source
1..3277
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/cissue_type="xylem"
1..3277
/gene="CesA2"
21..3119
/gene="CesA2"
/codon_start=1
/product="cellulose synthase"

1892 ACAGTCTTCTTTCGACATCAACATGAAGGGGCTCGACGGGATCCAAAGCCCGGTGTACGTC 1951
1776 ACTGTTTCTTCGATATAAATCAATGAAGGCTAGATGAATTCAGGGTCCAGTGTATGTG 1835
1952 GGGACAGGGTGCCTGTTTCGGCGCCAGGCGCTCTACGGCTACAAACCTCCCAAGGACCC 2011
1836 GGCACAGGATGCGCTTTTCAAAAGGCAAGCTTTGTATGGCTATGACCTCCCAAGGATCCA 1895
2012 AAGAGGCCAAGATGGTACCTCGACTGCTGCGCCGCTCTCGGCGCGCAAGAGCGGAAA 2071
1896 AAGGCCCAAGATGAACCTCGGACTGCTGCCCATGTTTGGAGCTGCCTCAAAAGAG 1955
2072 CAGCCAAAGGACGGGCTGCGGAGGCGACCGCTGATATGGGAGTAGACGCAAGGAG 2131
1956 AATGCTAAGACTGGTGCAGTTGTAGAAGGAATGGATA-----ATAATGACAAGGAG 2006
2132 ATGCTCATGTCCACATGAATTCGAGAGCGGTTGGGCGAGTCCGGGGGTTTCGTCACG 2191
2007 CTGTTGATGTCACATGAATTTTGAAGGAAGTTTGGACAATCAGCAATTTTTCGTAAC 2066
2192 TCGACGCTGATGAGGAAGCGGCTCCCTCTTCGTCGAGCCCGCGCGCTCCCAAG 2251
2067 TCAACTTTAATGAGAGAGTGTGTACCTCTTCCTCGAGTCCGGAGCTCTGCTAAAG 2126
2252 GAGGCCATCCATGTCTACGTGCGGCTACGAGGACAAAGCCGACTGGGGGCTGGAGCTG 2311
2127 GAAGCCATCATGTGATCAGTTGTGGATATGAAGACAAATCTGAATGGGAGCTCGAGCTG 2186
2312 GGTGTGATCTACGGGTGATCAGGAGGACATCTGACGGGGTTCAAGATGCACTGCGGC 2371
2187 GGTGTGATTTACGGTTTCGATCAGGAGGATATTTCTGACAGGTTTAAAGATGCAATTCGT 2246
2372 GGTGTGCGCTCCGTGACTCATGCGGAGCGGCGCGGTTCAAGGGCTCGGCGCGGATC 2431
2247 GGTGTGAGGTCTAATTACTGTATGCCAAGAGAGCTGCATTTAAGGGTTGAGTCCCATC 2306
2432 AATCTATCGGACCGCTCTCAACACGAGTCTCCGTTGCGGCGCTCGGCTCGAGATCTTC 2491
2307 AATCTATCAGATCGGCTAAACCAAGTGTCCGCTGGGCTCTTGATCTGTTGAAATTTT 2366
2492 TTCAGCGGCACAGCCCCCTGCTGTACGGCTACAAGACGGCAACCTCAAGTGGCTGGAG 2551
2367 TTCAGTGGTCACAGCCCTTAATTGGTATGGCTACAAGAAAGAAAGTCAAGTGGCTCGAG 2426
2552 CGCTTGGCTTACATCAACACACATCTACCCCTTCACCTCGCTCCGCTGCTCGGCTAC 2611
2427 AGGTTTGGGTATGTGAACACAACTAATCTACCCCTTACCTCTTAGCACTCGTTGATAC 2486
2612 TGCAACCTCCCGCGCTGTGCTCCTCACCGGCAAGTTTCATCATATGCGGTGATTAACG 2671
2487 TGTGCTCCTCGCATCTGCTGCTTACTGTATAAATTTATCATGCGCGAGATAAGCACC 2546
2672 TTCGCGAGCTCTTCTTCATGCGGCTTTCATGTCTCATCTTCGCGAGGCGATCCTGGAG 2731
2547 TTTGCAAGTCTTTTCTCATTTGCTGTTTGTCTCAATCTTTTCCAGCGGCATCTTGAG 2606
2732 ATCGGTGGAGCGGGTGAGCATCGAGGAGTGTGGAGGAACAGAGCTTCTGGGTATC 2791
2607 CTCAGATGGAGCGGAGTAGCAATGGAGATGGTGGAGAAACGACGCAATTTCTGGGTATA 2666
2792 GCGCGGTGTCCGCGCATCTTTTCGCGGTGCTGCAGGGCTGCTCAAGGTCTTCGCGCGG 2851
2667 GGTGGTGTGCTGCTCACCTCTTGTGTTGTTTCCAAAGTCTTCTGAAAGTTTGTAGCAGGT 2726
2852 ATGACACCAACATTCACGTCACCTCCAGGCGCACCGGCGACGAGACGAGATTCGCC 2911
2727 ATCGACTAAATTTCACTGTGCATCCAGGCTAC-----AGACGATGACGATTTTGA 2780
2912 GAGCTCTACGCTTCAAGTGGACACAGCTCTCTCATCCCGGCCACACGCTGCTCATCAT 2971
2781 GAGCTTTATGCCTTTAAATGGAACAACTCTGCTTATCCCTCAACCACTATCTTANTCATC 2840

2972 AAGCTCATCGGCGTGTGTCGCCCGGCTATCTCCGAGCGCATCAACACGCGGTACCAAGTCTCTGG 3031
2841 AACCTTGTGTGAGTTGTGCTCGAGTCTCAGATGCATAAACAAATGGGTACCAAGTCTATGG 2900
3032 GGGCCCCCTTCGGCAAGCTCTTTCGCTTCTGGGTCATCGTCCACCTCTACCCGTTTC 3091
2901 GGACCTCTAATTGGGAAGCTCTTCTTTCGCTTCTGGGTGATTGTCTCATCTCTACCCATTC 2960
3092 CTCAAGGGGCTCATGGGCGCCAGAACAGGAGCGCCACCGTGTGTTTGTCTCATCTGGTCCATT 3151
2961 CTCAAAGGCTAATGGGAGGCAAAACAGACACCGACTATTGTGTTATATGTCAGTG 3020
3152 CTGCTGGCTCCATCTTCTCCCTCTCTGGGTGAGGATCGACCCCTTTCATCTGTCAGGACC 3211
3021 CTCTGGCTTCCATCTTCTCTCTGCTTGGGTCCGATTGATCCATTGTGATGAAAAACC 3080
3212 AAGGCCCGGAGCTCAGGCAAGTGTGCGCATCAATTCCTGA 3250
3081 AGGGGACCTGACACCAAGCAATGTGGACTCAACTGTTAA 3119

RESULT 6
AY372244
LOCUS
DEFINITION
Zea mays cellulose synthase catalytic subunit 10 (Cesa10) mRNA,
complete cds.
ACCESSION
AY372244
VERSION
AY372244.1 GI:38532099
KEYWORDS
Zea mays
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 3470)
AUTHORS
Dhugga,K.S., Barreiro,R., Appenzeller,L., Wang,H., Niu,X.,
Carrigan,L. and Tomes,D.
TITLE
Cellulose formation and its role in determining stalk strength in
maize
JOURNAL
Unpublished
2 (bases 1 to 3470)
AUTHORS
Dhugga,K.S.
TITLE
Submitted (21-AUG-2003) Trait and Technology, Pioneer Hi-Bred
International, Inc., a DuPont company, 7300 NW 62nd Avenue,
Johnston, IA 50131, USA
FEATURES
Location/Qualifiers
1..3470
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
1..3470
/gene="Cesa10"
29..3265
/gene="Cesa10"
/note="Cesa10; ZmCesa10"
/codon_start=1
/product="cellulose synthase catalytic subunit 10"
/protein_id="AAR23310.1"
/db_xref="GI:38532100"
/translation="MDAGSVTGLAAGSHMRDELHVMRAREEPNAKVRADVKTCTVRC
ADEVGTREDGQPFVACAECPVPCPYEYSESGTQCPCQNTYKROKQCPRVBGD
EEGEPMDDFDEFPAKSPKPEHPVAFVYSENGEHPAQKWTGQTLSSTFGTSVAG
KDLAEAREMEGMEWKDRIDKWTQKQKRLNHDSDDDDDKNEDEYMLLAARAPL
WKVPTPSMINPYRIVILRLVLCFLKPRITTPATDAVLWLASVLCELWFAFSW
ILDQLPKAPVTRTYLDRALRYDREGACRLSPIDFFVSTVDPLKEPILITNTVL
SILAVDYPVDRVSCYVDDGASMLLFDALSETAEFAARWVPCKKFAVFPRAPEFVFS
QKIDLKDKVQPTFKERRAMREYEFKVRINALVAKQKPEEGVMQDQFPFGN
NTRDPHGMTQVYLGNGALDVEHELPLVYVSRKPGYNHHKKAAGMMNLVRVSAY
LTNAPFILNLDCHDYNNNSKAVREAMFLMDQLGKLCYVQPPQDFDGLDRLDRVAN
RNVPFEDINMGKLDGIQGPVYVYGCVFNRQALYGYDPPRPEKRPMTCDKWPSCCC
CCCGGGGKGAKDKDGGEPBRGLLGFKYKSKDKLGGSGVAGSKGGGLYKK
HQRAFELEIEEGLEGYDELESSLMSQSKSEKRFQSQSPVFIASTLVEDGGLPQGA

DPALIKAEATHVLSGVEEKEWKGSEIGWLVGSVTEIDLTGFKWHCHGKSVYCTPR
PAFKGSPINLSDELHVLRLWGLSVSEIFMSRCHPLRYAYGRLKWLNERPAYNTIYVI
PFTSIPILACTTIPAVCLLTGKFIPTLNLSIAFWIALFSLIATISVLSRWGVSIV
EDWRNEQFVWIGVSAHLFAVFGQFLKVLGVSDTFTVTSKAGDADAFDGLYLFFK
WTTLLVPTTLIIINMVGIVAGVSDAVNNGYSGWGPLFGKLFPSFWIVHLVPLKGL
MQRNRTPTIIVLWLSILLASIFSLVWVRIDPFIKPKAGPILKPGVBC"

ORIGIN

Query Match 45.4%; Score 1563.4; DB 8; Length 3470;
Best Local Similarity 70.9%; Pred. No. 2.8e-136;
Matches 2369; Conservative 0; Mismatches 671; Indels 303; Gaps 10;
QY 97 GGCAGAGCGCGGGTGTGCGCGCTGCTGCAACCGGAGACGAGCTGTGATCCGGG 156
DB 43 GGTCAACCGTGGCTCCGCGGGCTCGACATCGGGAGAGCTGATGTGATGCGCGC 102
QY 157 CCAGGAGACCCCAAGCC---GTCGCGGGGCTGAGCGGGCAGGTGTGCGAGATATCGG 213
DB 103 CCGGAGGAGCCGAAACCGCAAGGTCCGGAGCGCGACGTGAAGACGTGCCGCTGTGCGC 162
QY 214 CGAGAGGTGCGGTCTACCGTGGACGCGACCTTCTGTCCTGCAACGAGTGCCTT 273
DB 163 CGACGAGTGGGACGCGGGAGGACGCGGACGCTTCTGTCGCTGCGCGAGTGCCTT 222
QY 274 CCGCTGTGCGGCTCTACGAGTACGAGCGCGGAGGCGACGAGTGCCTCCCA 333
DB 223 CCGCTGTGCGGCTGCTACGAGTACGAGCGCGGAGGCGACGAGTGTGCTGCCGCA 282
QY 334 GTGCAAGACGCTTACAGCGCTTCAAGGGAGCGCGAGGTTCCGCGGACGATGACGA 393
DB 283 GTGCAACCGCTTACAGGCGCCGAAAGGTGCGGAGGTGAAAGGAGCGAGGAGGA 342
QY 394 GG---AGGACATCGACGACCTGGAGCAGCGATTCACATCGACGACGAGATCAGCAGAG 450
DB 343 GGGCCCGAGATGAGACGACTTCGAGGACGAGTTCGCGCCCAAGAGCCCAAGAGCCTCA 402
QY 451 GCAGCTGGAGGCGAATGACGACGACGCGAGTACGAGCGGATGTCGACGCGCAGAT 510
DB 403 CGAGCTGTGCGGTTCGAGTCTACTCGAGAACGCGGCGACCGCGCGCAGAAATGGCG 462
QY 511 GAGCTACGGAGGGGCGCCGACGACGCGGCGCAACACACCCCGCAGATCCGCGCCAT 570
DB 463 GACGGTGG-----CAGAGCTGTCTCTC 486
QY 571 CATCACCGTCCCGTCCGTGCGGTGAGCGGTGAGTTTCCGATTACCAACGCGGTATGG 630
DB 487 CTTTACCGGAAGCGTCGCGGGAAG----- 512
QY 631 CCAGGCGAGGTCTCGTCTTCCCTGCGACAAAGCGATCCATCCGTGTGTGTGAGCC 690
DB 513 -----ACCTGGAGGCGGAGAG 528
QY 691 AGGAGTGCCTAAGTGGGACGAGAAAGATGAGCTGGAAGGAGGAGTGAACGACTG 750
DB 529 GAGATGAGGGGAGCATGAGTGAAGGACCGGATCGACAGTGAAGACCAAGACGGA 588
QY 751 GAAAGTCAAGCAGGCGATCTCGCGCGCGCGGATCCGGAAGACATGAGCAGCGCGAGT 810
DB 589 GAAGAGGGCAAGCTCAACACGACGACGAGCAGCGACGACGACAAAGAACGAAGCA 648
QY 811 G---GCACTGAACGACGAGGAGGAGCGCGTGTGAGGAAGGTGCGTCCGTCGAG 867
DB 649 GTACATGCTGTTGCCGAGGCGGACAGCGCGTGTGCGCGAAGGTTCCGATCCCGTCGAG 708
QY 868 CAAGGTGAACCGTACCGGATGATGATCGGTGTCGTCGTGTCGCTCCGCTTCTTCCT 927
DB 709 CATGATCAACCGTACCGATTCGTCATCGTCTCCGCTGCGTGTGCTGTCTGCTTCTTCT 768
QY 928 CCGGTACCGTATCCTGCAACCCCGTCCGCGACGCGCATCGGCTGTGGCTCGTCTCCATCAT 987
DB 769 CAAGTTCGCGATCAGCAGCGCGCCGACGAGCGCGTCTGTGTGCTGGCTCGGTCTCAT 828
QY 988 CTGGAGATCTGGTTCGCCATCTCTCGGATCTCTCGACAGTTCGCCAAGTGGTTCGCCAT 1047

DB 829 CTGGAGCTCTGGTTCGCTTCTCTGGATCTCTGGAACAGCTGCCAAGTGGGCGCGCT 888
QY 1048 CGACCGCGAGACGTACCTCGACCGCTCTCCCTCAGGTACGAGAGGAAAGGAGCGGCTC 1107
DB 889 GACCGGGAGACGTACCTGGACCGCTGCGGTACGACCGTGGGCGGAGGCGTG 948
QY 1108 GCTGTGTGCGCGGTGACCTGTGTGAGACGCTGAGACCGCTCAAGGAGCGCGCT 1167
DB 949 CCGGTGTCTCCCATCGACTTCTTCTGTCAGACGCTGAGACCGCTCAAGGAGCGCGCAT 1008
QY 1168 GGTGACCGCAACACCGTGTCTCCATCTCGCGGTAGATACCCCTGCGACAGGTCTC 1227
DB 1009 CATCACCGCAACACCGTGTCTCCATCTCGCGCTGACTACCCCTGGACCGCGTCAG 1068
QY 1228 CTGCTAGCTCTCCGACGACGCGCGCTGATGCTGACGTTTCGAGTCTGCTGCGAGAGGCG 1287
DB 1069 CTGCTAGCTCTCCGACGACGCGCGCTCATGCTCTTTCGACGCGCTGTCGAGACGCG 1128
QY 1288 CGAGTTCGCGCGCAAGTGGGTGCGCTTCTGCAAGAGTTTCGGAATCGAGCGCGCGCC 1347
DB 1129 CGAGTTCGCGCGCGCTGGGTGCGCTTCTGCAAGAGTTTCGCGTGGAGCGCGCGCC 1188
QY 1348 GAGTTCCTACTCTCTCAAGTTCGACTACTCAAGGACAAAGTGCAGSCCCACCTTCGT 1407
DB 1189 GAGTTCCTACTCTCGGAGAAAGTCACTACCTCAAGGACAAAGTGCAGCGCGCTTCGT 1248
QY 1408 GCAGGACGCGCGCATGAAAGAGAGTATGAGGAGTTCAAGTTCGCGATCAACGCGCT 1467
DB 1249 CAAGGAGCGCGCGCATGAAAGGAGTACGAGGAGTTCAAGTTCGCGATCAACGCGCT 1308
QY 1468 GTGGGCAAGGCGCATGAAAGTTCGCGCAGAGGCTGATCATGAAGACGCGACGCGTG 1527
DB 1309 GGTGGGCAAGGCGCAGAAAGGCGCGAGGAGGCTGCGTATGTCAGACGACGCGCGTG 1368
QY 1528 GCCCGGGAACAACACCGCGACCCCGCATGATCCAGTTCAGGTTCCTGGGCGACAGCGG 1587
DB 1369 GCCCGGGAACAACACCGCGACCCCGCGGTATGATCCAGTTCCTCGCAACGAGG 1428
QY 1588 CGCGCCACGACACCGAGGCGAACGAGTTCGCCGCTGCTGATGCTTCCGTCGAGAGCG 1647
DB 1429 CGCGTGGAGTGGAGGCGCACGAGTTCGCGCGCTCTCTAGTGTCCGTCGAGAGCG 1488
QY 1648 CCGCGATTCAGACCAACAAGAGCGCGCGCATGAACGCTCTGATTCGCGTCTCCG 1707
DB 1489 CCGCGGTAACAACCAACAAGAGCGCGCGCATGAACGCTGCTGTGCGCTCTCCG 1548
QY 1708 GGTGTGACCAACGCGCATTCATGCTCACTTGGATGTGATCTACATCAACACAG 1767
DB 1549 GGTGTCAACACGCGCGCTTCATCTCACTCGACTGCGACCTACGTCACCAACAG 1608
QY 1768 CAAGGCGCATCCGCGAGGCGCATGTGCTTCTCATGGACCTCAGTCCGCGGAGAGTCTG 1827
DB 1609 CAAGGCGGTGCGCGAGGCGCATGTGCTTCTCATGGACCGCGAGCTGGGGAAGAGCTCTG 1668
QY 1828 CTAGTTCAGTTCGCGCAGAGTTTCGACGCGCATTCGAGTTCGACGACGATACGTAACAG 1887
DB 1669 CTAGTTCAGTTCGCGCAGGCTTCGATGGCATTCGATTCGCGACGACGATACGCAACCG 1728
QY 1888 GAACACGCTTCTTTCGACATCAACATGAAGGGCTGGAAGGCGCATCCAGGCGCGGTGA 1947
DB 1729 CAACGCTGTCTTTCGACATCAACATGAAGGGCTGGAAGGCGCATCCAGGCGCGGTGA 1788
QY 1948 GGTGGGACAGGCTGTGTTTCGCGCGCAGGCGCTCTAGCGTTCACAACTCCCAAGG 2007
DB 1789 GGTGGGACAGGCTGTGTTTCGCGCGGAGGCTGTACGCGTTCAGCGCGCGCGCC 1848
QY 2008 ACCCAAGAGCGCCCAAGATGTGAC-----TGCGACTGCTGCCGCTG 2049
DB 1849 CGAAGAGCGGCCCAAGATGACGTGCGACTGCTGCGCGCTGCTGTGCTGCTGCTGCTG 1908
QY 2050 CTTTCGCGCGCAAGAGG----- 2067

Db 1909 CTTGCGCGGGCAAGCGCGGCAAGGCGCGCAAGGCAAGAGGGCGACGCGGGCGAGGA 1968
QY 2068 -----GAAACAGCCCAAGAGCGG 2085
Db 1969 GCGCGCGCGGGCGCTGCTCGGCTTCTACAGGAAGCGGAGCAAGAGCAAGCTCGCGG 2028
QY 2086 GCTGCCGAGGGGACCGCTGATATATGGGAGTAGATAGCGACAAGAG-----2131
Db 2029 CGGGTCGGTGGCGGAGCAAGAGGGCGCGGGCTGTATACAGGAAGCACAGCGCGCGTT 2088
QY 2132 -----ATGCT 2136
Db 2089 CGAGCTGGAGGAGATCGAGAGGGGCTGGAGGGGTACGAGAGCTGGAGGGCTCTCGCT 2148
QY 2137 CATGTCCCATGAACTTCGAGAAGCGGTTTCGGGCAGTCCGGCGGCTTCGTCACTGCGAC 2196
Db 2149 CATGTCGAGAAGAGCTTCGAGAAGCGGTTTCGGCCAGTCCGCGGTTCATCGCCTCCAC 2208
QY 2197 GCTGATGGAGGAAGGCGGCGTCCCTC-----CTTCGTGAGCGCCGCGCGCTCCTCAA 2250
Db 2209 GCTCGTCGAGGACGCGCGGCTGCGCAGGGCGCGCGCGCGAGCCCGCGCGCTCATCAA 2268
QY 2251 GGAGGCCATCCATGTTCATCAGCTGCGGCTACGAGGACACAGCGACTGGGGCTGGAGCT 2310
Db 2269 GGAGGCCATCCAGTTCATCAGCTGCGGATACGAGGAGAACCGAGTGGGGCAAGAGAT 2328
QY 2311 GGGGTGGATCTAGGGTGCATCAGGAGGACATCTTCAGCGGGTTTCAAGTGCATTCGCG 2370
Db 2329 TGGGTGGATCTATGGGTGCTGACAGAGGATATCTTCAGCGGGTTTCAAGATGCATTCGCG 2388
QY 2371 CGGGTGGCGCTCGGTACTGATGCGAAGCGGGCGGGTTTCAAGGGTTCGGCGCGAT 2430
Db 2389 GGGGTGGAAGTCCGTACTGACACGCGACACGGCGGGGTTTCAAGGGTTCGGCGCGCAT 2448
QY 2431 CAATCTATCGGACCGTCTCAACAGAGTGTCCGCTGGCGGCTGGGGTCCGTGAGATCTT 2490
Db 2449 CAACCTGTCTGATCGTCTCACAGGTGTTCGCGCTGGGGCTGGGGTCCGTGAGATCTT 2508
QY 2491 CTTTCAGCGCGCACAGCCCGCTGTGTACGCGCTACAGAGACGGCAACCTCAAGTGGCTGA 2550
Db 2509 CATGAGCCGCACCTGCGCGCTCGGTACGCTAC---GGCGCGCGCTCAAGTGGCTGA 2565
QY 2551 GGGCTTGGCTCATCAACACACACATCTACCCCTTTCAGCTTCGCTCCCGCTGCTCGCCTA 2610
Db 2566 GGGCTTGGCTCATCAACACACACATCTGATACCCCTTTCAGCTTCATCCCGCTCTCGCCTA 2625
QY 2611 CTGACCCCTCCCGCGCTGCTGCTCTCAACCGCAAGTTTCATCATGCGCTCGATTAGCAC 2670
Db 2626 CTGCACATCCCGCGCTGCTGCTCTCAACCGCAAGTTTCATCATCCACCGCTGAACAA 2685
QY 2671 GTTCGCAGCGCTTCTTCATCGCCCTCTTCATGTCCATCTTCGCGACGGGCATCCTGA 2730
Db 2686 CTTGCGCAGCATCTGGTTTCATCGCGCTCTTCCTGTCATCATCGACGAGCGTCTGA 2745
QY 2731 GATCGGTTGAGCGGGTGAGCATCGAGAGTGGTGAGGAACGACAGTTCTTGGGTAT 2790
Db 2746 GCTGCGGTGAGCGGGTGAGCATCGAGGACTGGTGGCGCAACGACAGTTCTTGGGTCTAT 2805
QY 2791 CGGCGCGGTGTCGCGCATCTCTTCGCGCTTCGTGCAAGGCGCTGCTCAAGGTCTCTCGCGG 2850
Db 2806 CGGCGCGGTGTCGCGCATCTCTTCGCGGTTCAGGGGCTTCCTCAAGGTCTTGGCGG 2865
QY 2851 GATCGACACCACTTCACCGTCACTTCAAGGCCACCGCGACGAGGACGACGAGTTTCG 2910
Db 2866 CGTGGACACAGCTTCACCGTCACTTCAAGCGCGCGGCGAGCGCGCTTCG 2925
QY 2911 CGAGCTCTAGGCTTCAAGTGAGACACGCTCTCATTCGCGCCACACCGCTGCTCATAT 2970
Db 2926 GGACCTCTACCTTTCAGGTGGACCACTCTGCTGGTGGCGCCGCCACACCGCTCATCAT 2985
QY 2971 TAAAGTCATCGGCGTGGCGGATCTTCGAGCGCATCAACAAAGGGTACAGTCCG 3030
Db 2986 CAACATGGTGGGATCTGGCGCGGCTGTCCGACGCGCGTCAACACGGCTACGGCTCG 3045

QY 3031 GGGGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTTCATGCTCCACCTCTACCCGTT 3090
Db 3046 GGGCGCGCTTCGGCAAGCTCTTCTTCCTTCTGGGTTCATGCTCCACCTCTACCCGTT 3105
QY 3091 CTTCAAGGGGCTCATGGGCGCCAGAACAGGACGCCACCGTTGTGTCTATCTGTGTCAT 3150
Db 3106 CTTCAAGGGGCTCATGGGAGGAGAACCGGACGCCACCATCTGTGTCTGTGGTCCAT 3165
QY 3151 TCTGCTGGCTCTCATCTTCTCCCTGCTCTGGGTGAGGATCGACCCCTTTCATGTCAGGAC 3210
Db 3166 CTTCTCTGGCTCCATCTTCTGCTGTGGGTGAGGATCGACCCCGTTTATCCGAGGC 3225
QY 3211 CAAGGCGCCGAGCGTCAAGGACAGTGTGGCATCAATTGCTGAGCT 3253
Db 3226 CAAGGCGCCCATCTCAAGCCATCGGAGTCGAGTCTGAGCT 3268

BT013780 3414 bp mRNA linear PLN 11-MAY-2004
Lycopersicon esculentum clone 132670R, mRNA sequence.
DEFINITION
ACCESSION
BT013780
VERSION
BT013780.1 GI:47105195
KEYWORDS
FLI_CDNA.
SOURCE
Lycopersicon esculentum (tomato)
ORGANISM
Lycopersicon esculentum
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 3414)
Kirkness, E.F., Wang, W. and Vazeille, A.
AUTHORS
Direct Submission
TITLE
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
JOURNAL
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1..3414
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132670R"
/tissue_type="fruit"
/note="TZXY89"

ORIGIN
Query Match 45.1%; Score 1552.6; DB 8; Length 3414;
Best Local Similarity 69.7%; Pred. No. 7.3e-195;
Matches 2209; Conservative 0; Mismatches 909; Indels 51; Gaps 6;
QY 89 GCGATGGAGGCGAGCGCGGCTGGTGGCGGCTCGCACAAACGGAACGAGTGTGCTG 148
Db 63 GCGATGGAGGCTAGTGCAGGCTAGTGTGCTGCTTACATAACCGGAATGAACCTTGTGCTC 122
QY 149 ATCCGGGGCCAGAGGACCCCAAGCGCTCGCGGCGCTGAGCGGCGAGTGTGCGAGATA 208
Db 123 ATTCATGGCCATGAAGAGCATAAAGCCATTGAAGGATTTGAGTGGACAAGTTGTGACATA 182
QY 209 TCGCGCGACAGGCTCGGGCTCACGGTGGACGGCGACCTCTTCGTCCTGCAACGAGTGC 268
Db 183 TGTGGGATGAATAGGCTTAACAGTGGATGTGATCTTTTGTGGCTTGCATGAGTGT 242
QY 269 GGCCTTCCCGTGTGCGGCCCTGCTTACGAGTACGAGCGCCGAGGAGGCGACGCGAAGCTGC 328
Db 243 GGTTCCTCTGTGTGCGGCTTGTACGAGTATGAGAGGAAGGAAGAACTCAACAGTGT 302
QY 329 CCCAGTGCAGACGCGCTTACAGCGCTCAAGGGGAGCCCGAGGGTTCGCGGGACGAT 388
Db 303 CCACAGTGCAGACGAGATACAAACGCTCTTAAAGGAGTCCAAAGGGTGGCGGAGATGAT 362
QY 389 GACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAACATCGACGAGAGATCAGCAG 448
Db 363 GATGAGGAGATATTGACGATATTGACATGATTAACAGTTGATGATGA-----412

Db	2547	CTCCTTGCCTACTGCACTCTCCCTGCTGCTGTTTCTCACTGGAATAATCAATGCCA	2606
QY	2660	TGCAATTAGCAGGTTCCGAGCCTCTTCTTCAATCGCCCTCTTCAATGTCATCTTCGCGAGC	2719
Db	2607	GAGATAAGCAGGTTAGCAGTCTGTTTTCATTTGCTCTTCTCTCTATCTTCAACAGC	2666
QY	2720	GGCATCTCGAGATCGGTGGAGCGGGTGAGCATCGAGAGTGGTGGAGGAACGAGCAG	2779
Db	2667	GGAAATCTTGGCTAAGTGGAGTGAAGCATCGAGGAATGGTGGAGGAATGAACA	2726
QY	2780	TTCTGGGTCTATCGCGCGGTGTCGGCGCATCTTTCGCGGTGCTGAGGCGCTCTCAAG	2839
Db	2727	TTTGGGTATTGGAGGTGTATCTGCCACCTTTTTCGTGTAGTACAAGGTCTTCTCAAG	2786
QY	2840	GTCTCGCGGGATCACACCACTTTCACGTCACCTCAAGGCCACCGCGACGAGGAC	2899
Db	2787	ATTTCGCTGGCATAGTACCACTTTTCTGTGACATCTTAGGCAACAGATGATGAGGAC	2846
QY	2900	GACGAGTTCGCGAGCTTACGCTTCAAGTGAGCACACGCTCTCTATCCCGGCCACG	2959
Db	2847	-----TTGCGAGGTATATGCTTCAAGTGAGCTACTCTCTTATTCCTCAACA	2900
QY	2960	CTGCTCATATTAACCTCATCGGCTCGTGGCGGCATCTCGAGCCATCAACACGGG	3019
Db	2901	ATTCTCATATCAACTTGGTGGGTGTAGCCGGAATCTCTGACCTATCAACACGGA	2960
QY	3020	TACCACTCTCGGGGCCCTCTTCGGCAAGCTCTTCTCGCCTTCTGGGTCTATCGTCCAC	3079
Db	2961	TACAAATCATGGGTCCCTTTTCGGAAGCTTTTTCGCTTTGGGTCTATGTTCAAT	3020
QY	3080	CTCTACCCCTTCTCAAGGGGCTCATGGGGCGCCAGAACAGGACGCCACCGTGTGTC	3139
Db	3021	TTATATCACTTCTCAAGGCTCTTATGGCGAGCAGAAACAGGACTCCAACCTTGTGTA	3080
QY	3140	ATCTGGTCACTTCTGCTGCTCCATCTTCTCCCTCTCTGGTCAAGTCAACCTTTC	3199
Db	3081	ATATGCTCTATCTTTTAGCCTCCATCTTTTCCTGCTCTGGGTTCGAATGATCACTTT	3140
QY	3200	ATCGTCAGGACCAAGGCCCGGACGTCAGGAGTGGCATCAATGCT	3248
Db	3141	GTGCTCAAAACCAAGGACCATGTCAGGATGTCAGTGAATGCT	3189
RESULT 8			
LOCUS	AY139754	3355 bp	mRNA linear PLN 02-SBP-2002
DEFINITION	Arabidopsis thaliana AT5g17420/T10B6_80 mRNA, complete cds.		
ACCESSION	AY139754.1		
VERSION	AY139754.1 GI:22654964		
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 3355)		
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Huan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.		
TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3355)		
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Huan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.		

TITLE		Submitted (09-AUG-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
The Salk, Stanford, PCEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H., Cheuk, R., Kim, C. J., Bowser, L., Chan, M. M., Chang, C. M., Dale, J. M., Deng, J. M., Hsuan, V. W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Palm, C. J., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Wallender, E. K., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Davis, R. W., Theologis, A., and Ecker, J. R.			
Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.			
FEATURES	Location/Qualifiers		
source	1..3355		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="5"		
	/clone="RAFL09-35-F05 (R22199)"		
	/ecotype="Columbia"		
	46..3126		
	/note="cellulose synthase catalytic subunit (IRX3)"		
CDS	/codon_start=1		
	/product="AT5g17420/T10B6_80"		
	/protein_id="AA098075.1"		
	/db_xref="GI:22654965"		
	/translation="MEASAGLVAGSHNRNELVVIHNEBPKLNLDGQFCBICGDOI GLTVEGDLFVACNECGFPACRCEYERECTQPCQCKTRYKRLGRSFRVSGDEDEB DIDDIEYEFNIEHQDKKHSAMLYGKMSYGRGDEDDENGFRFPVIAAGHSGEFPV GGGYNGEHLGHRVHPYPSSSAGSEGGRRERMDWKQHNGLGDEDDPENGGLIDE ARQPLSRKVPYASSINPYRMVIVARLVILAVLRNLNPHVLDALGLWLTSTVCSIM FAVSWILAMDYVPEKISCYVSDDCASMLTPESLSETEAFARKWPKFCCKGSIERAP SNTVLSILAMDYVPEKISCYVSDDCASMLTPESLSETEAFARKWPKFCCKGSIERAP EMYFTLVLDQKVHPTFVKERRAKREYEEFKVIRNAQVAKASKVPLEGHIIMODGT FPMGNNTKDHGMIQVFLHSGGDFVEGHELRPLVYVREKRPQGHKKKAGAMNALV RVAGVLTNAPMLNLDCDHYVNSKAVREAMCFPLMDPOIGKVCYVQFPQRFEDGDTN DRYANRNTVFDINNKGLDGIQGVYVGTGCKRQALYGYEPKGRPKRMISGCCC PCGRRRNKKPQKNDMNGDVAALGAGBDKHLMSMNFEKTFQSSIFVSTLMEEE GGVPPSSPAVLLKEAIVHISGVEDKTEWGLQWYGSITFDILTGFPMKCRGWRSS IYCMKPRAPFKGSAPINSLRNLQVLRWALGSVEIIFSRHSPLWYGYKGGKLVKWERF AYANTITPFTSIPLLAYCILPAICLLTDKDFIMPPISTPASLFFISLMSIIVTGILE ELWAGVSIIEWNRNQFWIGGISAHLFVAVQGLLKLAGIDTNFTVTSKATDDDDFG ELYAFKWTTLIPTTTLINIVGVAGISDANNNGYSWGLFGLKLPFSFWIVHLY PFLKGLMGRQNRTPITVIVWSVLLASIFSLMWLRIDPFLVLTGKPDTSKGCINC"		
ORIGIN			
Query Match	44.1%;	Score 1519.2;	DB 8; Length 3355;
Best Local Similarity	69.8%;	Pred. No. 1.8e-190;	
Matches 2213;	Conservative	0; Mismatches 863;	Indels 96; Gaps 8;
Qy	88	CGCGATGAGGCGCAGCGCGGCTGTGTGCGCGGCTCGCACACCGAACGAGCTGGTCT	147
Db	42	CCCTATGGAAGCTAGCGCGGCTTGTGCGCGGCTCTCATAAACCGTAATGAACCTAGTCGT	101
Qy	148	GATCGGGGGCCACGAGGACCCCAAGCGCTGCGGGCGCTGAGCGGGCAGGTGTCGAGAT	207
Db	102	CATTACACCACTGAAGACCAAGGCTCTGAAGAATCTAGATGGACAAATCTGTGAGAT	161
Qy	208	ATGCGGCGACGAGTTCGGGCTCACGCTGGACGCGACCTCTTCGTGCCTCGCAACGAGTG	267

QY 2419 GTCCGGCCGATCAATCTATCGACCGTCTCAACCGAGTGTCTCGGTGGCGCTGGGGTC 2478
Db 2301 TTCCAGTCTCTTAAATCTATCAGACAGGTTAAACACGAGTTTTCGTTGGGCATCTTGATC 2360
QY 2479 CGTCGAGATCTCTTCAGCGGCGACACGCCCTCTGTGTAGCGGTACAAGAACGGCACTT 2538
Db 2361 GGTGTAGATATTTTTTCAGCGGCGACAGTCTCTCTGTGTATGGTATCAAAAGGAGGCAAACT 2420
QY 2539 CAAAGTGGCTGGAGCGCTTCGCCCTACATCAACACACCATCTACCCCTTCACCTCGCTCCG 2598
Db 2421 CAAAGTGGCTGGAGCGTTTGGCTTATGCCACACACAACTACCCCTTCACATCTATACC 2480
QY 2599 GCTGCTCGCTATCTACACCCCTCCCGCGTCTGCTCTCCACCGCAAGTTCATCATGCC 2658
Db 2481 ACTTCTTGGCTACTGTATCTTCCAGCCATCTGCTCTTACTCTACGACAAATTCATCATGCC 2540
QY 2659 GTCGATTAGCACGTCGCGAGCGCTCTTCTCATCGCCCTCTTCTATCTCCATCTTCGCGAC 2718
Db 2541 ACCGATAAGCACATTTGCTAGTCTCTTCTCATCTCTACCTGTTATGTGATCATTTGTAAC 2600
QY 2719 GGGCATCTCTGGAGATCGGTGGAGCGGGTGGAGCATCGAGGAGTGGTGGAGGAACGAGCA 2778
Db 2601 GGGAACTCTTGGATTGAGATGGAGCGGAGTTAGCATTTGAGGAATGGTGGAGAAACGAGCA 2660
QY 2779 GTTCTGGGTATCGCGCGCGGTGTCGGGCGATCTCTTCGCGGTCTGTGAGGCGCTGCTCAA 2838
Db 2661 ATTCTGGGTCAATGGAGGAATCTCAGCTCATCTCTTTTGGGTTGTCAAAGTCTCTCTCAA 2720
QY 2839 GGTCTCGCGGATCGACACCAACTTCCAGTCACTCCACGCGCCATCCAGGCCACCGGACGAG 2898
Db 2721 AATCTTAGCAGGCAATTTGACACAAACTTCCAGTCACTCAAAAGGCAAC-----AGATGA 2774
QY 2899 CGACGAGTTCGCGAGTCTTACGCCCTTCAAGTGGACACAGCTCTCTCATCCGCGCCACCCAC 2958
Db 2775 TGAATGACTTTGGAGAACTTTACGATTTCAATGGACACAACTGCTGATCCCTCCACAAC 2834
QY 2959 GCTGCTCATCTAACTCACTCGCGTCACTCGCGCGGATCTCCGAGCCCATCAACACCG 3018
Db 2835 TGTCTTAAATCATAAACATTTGTCGGCTTGTTCAGGCGATCTCAGATGCCATTAACAATGG 2894
QY 3019 GTACCACTCTGGGGGCCCTCTTCGCGAGCTCTTTCGCGCTCTTGGGTCTATCTCTCA 3078
Db 2895 ATATCAGTCTTGGGAGCTCTATTTGGTAACTCTTCTCTCTCTTTTGGGTCTATTTCA 2954
QY 3079 CTTCTACCCGCTTCTCAAGGGGCTCATGGGGCGCCAGAACGAGCAGCCACCGCTTGTGT 3138
Db 2955 TCTCTACCCATCTCTCAAGGCTCTATGGGTGACAGACAGACACCAACCATTTGGT 3014
QY 3139 CATCTGGTCACTCTGCTGGCCCTCATCTTCTCCCTGCTCTGGGTGAGATCGACCCCTT 3198
Db 3015 GATTTGGTCACTGTTATTTGGCATCTATCTCTCTTGTGCTTGGGTGAAGATTCATCTTT 3074
QY 3199 CATCGTCAGACCAAGCGCCGACGCTCAGGCACTGGCATCAATTCGCTGA 3250
Db 3075 TGTGCTCAAGACCAAGGACCTGACACTTCCAAAGTGTGGCATCACTGCTGA 3126

RESULT 9
BT004543
LOCUS BT004543
DEFINITION Arabidopsis thaliana AT5g17420/t1086_80 gene, complete cds.
ACCESSION BT004543
VERSION BT004543.1
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.
1 (bases 1 to 3081)
REFERENCE
AUTHORS Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W.,
Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J.,

Lam B., Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M.,
Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
Southwick A., Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K.,
Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A. and
Ecker J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3081)
AUTHORS Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W.,
Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J.,
Lam B., Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M.,
Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
Southwick A., Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K.,
Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A. and
Ecker J.R.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2003) Salk Institute Genomic Analysis Laboratory
REFERENCE (SIGNAL), Plant Biology Laboratory, The Salk Institute for
AUTHORS Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.
The Salk, Stanford, PGE (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R.,
Chen, H., Kim, C.J., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M.,
Dale, J.M., Huan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and
Ecker, J.R.
Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
Location/Qualifiers
1..3081
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U22199"
/ecotype="Columbia"
/note="this clone is in pUNI 51"
1..3081
/note="cellulose synthase catalytic subunit (IRX3)"
/product="AT5g17420/T10B6_80"
/protein_id="AA042789.1"
/db_xref="GI:28416517"
/translation="MEASAGLVAGSHNRNELVVHNEPRLKLNLDGQFCBICGQOI
GLTVGDLFVACNECFACRYEYEREGTQNCQCKRYKRLGSRVGEDEEE
DIDDIYEENIEHQDKHSAEAMLYKMSYGRPEDENGRFPVIAAGHSGEPV
GGYNGEHLKRVHPYPSBAGSEGGWRMDWLQHLNGLPDDPDPFGIDITN
AQPLSRKVPKASKINPVMTVARLVLAFLRVLRNLPVHDALGLMLTSVCEITW
FASVILDMDFPKPIERTYDLRLSLRYEREGEPNMLAPDVVFVSTVPLKEPPLVT
SNVLSILADFPVEKISCTYSDGASMLTFESLSETAFARKWVPFCRKFSTPRAP
ENYFTLKVDYLDKVHPTFKERRARKREYEEFKVIRNAQKAKSVPLGEGTIMQDT
RVAGLVTAAPFNLNLDCHYVNSKAVREACFLMDPQKGVKQVQFQFPGIDITN
DSVANRVTFFDINMKLDGIOGPVYVGTGCVFKQALYGYEPKPKPKMISCCGC
PCFGRKNNKFSKNDMNGDVAALGAEAGDEKHELMSEMFETFGQSSIFVTSLMEE
GYCPSSSPAVLLKEAHSICGYEDKTEWTELGWIYISITLIDITLFGHCKWGRS
IYCMKPRKPAKSGAPINSLDRNLQNLWALGSEVFFSHSLPMLYKYGKGLKWLRF
AVANTTIVPTISPIILAYCILLTKDKTMTPTSTFASLFFISLFSIIVTGILE
LWMSGVISEWNRNQFVIGGISAHLFAVVOGLKILAGIDTNTFVTSKATDDDFG
ELYAKWTKLLIPTPTLVLIINIVGVVAGISDAINNGYQSGWGLFGKLFKFSFWIVLY
PFELGLMGRQNTPIVINSVLLASIFSLWLRIDPFLVLTGKPTDSKGCING"

ORIGIN

Query Match		44.1%; Score 1518.4; DB 8; Length 3081;
Best Local Similarity		69.8%; Pred. No. 2.3e-150;
Matches 2211; Conservative		0; Mismatches 861; Indels 96; Gaps 8;
Qy	92	ATGAGGCGCAGCGCGGCTGCTGCGCGGCTCGCACAAACCGGAGCGAGCTGCTGCTGTC 151
Db	1	ATGGAAGTAGCGCGGCTTGTGCGCGGTTCTCATACCGTAATAGTAACCTGCTGCTGCT 60
Qy	152	CGGGGCCACGAGGACCCCAAGCCGCTGCGGCGCTGAGCGGCGAGGTGTGCGAGATGCG 211
Db	61	CACAAACATGAAGACCAAGCCCTCTGAAGAATCTAGATGACAATCTGTGAGATATGT 120
Qy	212	GGCGACAGAGTGGGCTCAGGTCGACGGCGAGCCCTCTTCTGCGCTCGAACGAGTGGGCG 271
Db	121	GGAGATCAGATTTGGTTTAAACAGTAGAAGGAGACCTCTTCTGAGCTTGCATGAGTGGT 180
Qy	272	TTCCCGCTGTGCGGCGCTGTCTACGATACGAGCGCGGAGGCGACGACAGTGCCTCC 331
Db	181	TTTTCGCGGTGACACCTTGTCTATGATGACGAGAGAGAGAGAACACAAACTGTCTCT 240
Qy	332	CAGTGCAGAGCGGCTCAAGCGCTCAAGGGGAGCCCGAGGGTTGCCGGGACGATGAC 391
Db	241	CAGTGTAAAGCTCGTTACAGCGTCTCAGAGAAAGCCCAAGAGTGGAGGAGATGAAGAC 300
Qy	392	GAGGAGCATTCGACGACCTGGACGACGAGTTCAACATTCGACGACGAGATTCACGAGG 451
Db	301	GAAGAAGATATTGATGATATTGATGATGAATTTAATATCGAAATGAACAAAGA----- 353
Qy	452	CAGCTGGAGGCGACATGCGAGACAGCCAGATCACCGAGGCGATGCTGCACGGCAGGATG 511
Db	354	-----TAAGCATAAAGCATTTCTCTGAGGCTATGCTTTTATGGGAAATG 396
Qy	512	AGCTACGGGAGGGCCCGACGACGCGGACGCGCAACACACCCCGCAGATCCCGCCCATC 571
Db	397	AGCTATGGAAGAGTCTCTGAGATG-----ATGNAATGGGAGATTCCCACTGTT 447
Qy	572	ATCACCGGCTCCCGCTCGGTCGCGGTGAGCGGTGAGTTTCGATTAACAAGGATATGGC 631
Db	448	ATAGCTGGTGGTCA-----TAGTGGAGAAATTCAGTTGGAGGAGGTTATGGT 495
Qy	632	CAGGCGAGGTCTGCTTCCCTGCACAGGCGATCCATCCGTACCTGCTGTCTGAGCCA 691
Db	496	AATGGAGAAC-----ATGGGCTTCATAAGCGTGTGCACCATATCATCATCTGAAGCT 549
Qy	692	GGGAGTGCCAAAGTGGGACGAGAAAGAAAGTGAAGTGGAGGAGAGATGGAGCACTGG 751
Db	550	GGGAGT-----GAGGAGGATGGCGGGAAGAAAGATGGATGCTGG 588
Qy	752	AAGTCCAGAGGGCATCTCGGCGGCGGCGGATCCCGAAGACATGGAGCGGACGTG 811
Db	589	AAGCTCCAGCATGGAATCTTGGGCGCAGAACCCAGATGATC-----CTGAGATG 639
Qy	812	GCACTGAACGACGAGGCGGAGCGCGCTGTCTGAGGAGGTGTCGATCGCGTCCGAGCAAG 871
Db	640	GGACTGATGACGAGGACGCGACCCCACTCTCGCGAAAGTTCCCATTTGCTCAAGCAAG 699
Qy	872	GTGAACCCGTACCCGATGGTATCGTGTGCGTCTCGTTGTGCTGCTCTTCTCTCCGG 931
Db	700	ATCAATCCATATCGATGGTTCATCGTGTAGGCTTGTGATCTTAGCAGTTTTCGCGG 759
Qy	932	TACCGTATCTGCAACCCCGTCCGGACGCCATCGGGCTGTGGTCTGCTTCCATCATCTGC 991
Db	760	TATAGGCTCTGAATCCAGTGCATGATGCTCTGGGATATGGCTGACCTCTGTGATCTGT 819
Qy	992	GAGATCTGTTCCGCATCTCTGATCTCTGACCATGCTCGACCAAGTTCCTCAAGTGGTCCCATCGAC 1051
Db	820	GAATCTGGTTCGTGCTCTCTTGGATTTCTTGAATCTTGAATCAGTTTCCCAAGTGGTTCCTATTGAA 879
Qy	1052	CGCGAGACGTACCTTCGACCGGCTCTCCCTCAGGTACGAGAGGAGGAGGCGGCTGCTG 1111
Db	880	CGTGAGACCTATCTAGATCGGCTTCTCTCAGGTACGAGAGAGAGGATGAAACCAATATG 939
Qy	1112	CTGTGCGCGGTGACCTGTTCTGTGAGCACGGTGAACCGCTCAAGGAGCGCCCGCTGGTG 1171
Db	940	CTTGCCCTCTGTAGATGCTTTTGTGAGTACGGTGAACCCATTTGAAGGAGGCTTCCCTCTGC 999
Qy	1172	ACGCCCAACACCGTGTCTCATCTCGCGTAGACTACCCCGTGAACAAGGTCTCTCTGC 1231
Db	1000	ACATCCAACTACTGTGCTCAATCTTGCCCATGACTACCCAGTTGAGAAATCTCTCTGC 1059
Qy	1232	TACGTCTCCGACGACGCGCGCTCGATCTGACGTTTCGAGTTCGCTGTTCGGAGACGCCGAG 1291
Db	1060	TATGTCTCTGACGACGGTGTCAATGCTTACATTCGAATCTCTCTCGAAACTGCTGAG 1119
Qy	1292	TTCCGCGCAAGTGGGTGCGCTTCTGCAAGAGTTTCGGCATTCGAGCCCGCGCCCGGAG 1351
Db	1120	TTTGCAAGAAATGGGTTCCCTCTGTAAGAAATTTCCATAGAGCCACGCGGCAACCGGAG 1179
Qy	1352	TTCTACTTCTCGCTCAAGGTTCGACTACCTCAAGCAAGGTGACAGCCACCTTCGTGCGAG 1411
Db	1180	ATGTACTTTCAGTTGAAAGTTGATTTATCTCCAGGCAAAAGTCCACCCAACTTTGTTAAG 1239
Qy	1412	GAGCGCGCGCCATGAAGAGAGAGTATGAGGAGTTCAAGGTTCGGATCAACGCGCTGGTG 1471
Db	1240	GAACTGCGAGCCATGAAGAGAGAAATATGAGGAGTTCAAGGTTCAGGATCAACGCTCAAGTG 1299
Qy	1472	GCCAGGCCATGAAGGTCCGCGAGAGGGGTGATCATGAAGGACGCGACGCGCTGGCCC 1531
Db	1300	GCGAAGGCTTCAAGGTTCTCTAGAAAGTTGGATCATGCAAGATGGAACAACCGTGGCCA 1359
Qy	1532	GGGAAACAACACCCGCGACCCCGCATGATCCAGGTGTTCTTGGGCGCACAGCGGGCGG 1591
Db	1360	GGGAAACAACACCGAGGACCCCGGTATGATCAGTCTTCTTCGCGCACAGCGGAGGA 1419
Qy	1592	CACGACACCGAGGCGCAACGAGTCTCCCGCTCTGTGTAAGTCTCTCCGTGAGAACGCGCG 1651
Db	1420	TTTGATGTCGAGGGCATGAGCTTCTCGGCTGTGTGACGTGTCCTCGTGAAGACGCTCT 1479
Qy	1652	GGATTCCAGACACCAAGAGAGGCGGCGCATGAAACGCTCTGATTCGGCTCTCCCGCGTG 1711
Db	1480	GGTTTTTCAACACCAAGAAAGGTGGCGCCATGAATGCCCTGGTTTCGAGTGGCAGGCGTA 1539
Qy	1712	CTGACCAACGCGCATTCATGCTCAACTTGACCTGTGATACGTCTCTCCGTGAGAACGCAAG 1771
Db	1540	CTCAAAATGCTCTTTTCATGCTGAACTTGACCTGTGATCACTATGTTAAACAACAGCAAG 1599
Qy	1772	GCCATCCGGGAGGCGATGTGCTTCTCATGAGACCTCAGGTTCGCGCGGAGAGGTCTGTAC 1831
Db	1600	GCGTGAAGGAGCAATGTGTTTTTGTATGATCTCAGATGGAAGAGGTCTGTCTAT 1659
Qy	1832	GTTTCAGTTCGCGAGAGTTCGAGCGCATCGAGTGCACGACCGATAGCTTAACAGGAC 1891
Db	1660	GTTTCAGTTCCTCAAGGTTTGTATGGCATTTGACACAAACGATCGTTTACGCCAACAGAAAC 1719
Qy	1892	ACGCTCTTCTTCGACATCAACATGAAGGGCTGGAGCGCATCCAAAGCCCGGTGTACGTC 1951
Db	1720	ACAGTCTTCTTGTATCAATATGAAGGTCTAGATGGAATCCAAAGTTCAGTTTACGTT 1779
Qy	1952	GGGACAGGGTTCGTTTCGCGCGCGCAGGCGCTCTACGGGCTACACCCCTCCCAAGGAGCCC 2011
Db	1780	GGTACTGTTGTGTTTTTCAACGACAGCTCTGTATGTTTATGAACCAACAGAGGCTCT 1839
Qy	2012	AAGAGGCCCAAGATGGTGAACCTGGACTGTCGCCGCTGCTTCGGCCGCGAGAGCGGAA 2071
Db	1840	AAACGTCCAAAGATGATAAGCTGTGGTTGTGCTCTTGTCTTGGTGGCGCGGAGAAAGAT 1899
Qy	2072	CA-----CGCCAGGAGCGGCTGCCCGGAGGCGACCGCTGATATGGGAGTAGATAGC 2122
Db	1900	AAGAAATTTTCCAGAAATGACATGAATGGTGAAGTAGACGCTTGGAGGAGCAGAAAGT 1959
Qy	2123	GACAAGGAGATGCTCATGTCCCATGAACTTTCGAGAAAGCGTTTCGGGCGAGTCCCGCGG 2182
Db	1960	GATPAGGAGCATTTGATGCTGAAATGAACCTTTGAGAAACATTTGGGCAATCATCCATC 2019
Qy	2183	TTGTCACGTCGACGCTGATGGAGGAAGGCGGCTCCTCTCTGTCGAGCGCCCGCGCG 2242

Db 2020 TTTGTAACTCAACTTTGATGGAAGAGGTGGTGTCTCTCCGTCATCAAGTCTCGAGTG 2079
Qy 2243 CTCCTCAAGAGGCCATCCATGTCATCAGCTGGGCTACGAGCAAGACCGACTGGGG 2302
Db 2080 CTCCTTAAGAGCAATCCATGTCATAGCTCGGTATGAACAAGACTGAGTGGGA 2139
Qy 2303 CTGAGCTGGGTGGATCTACGGGTGATCACGGAGGACATCTCTGACGGGGTTCAAGATG 2362
Db 2140 ACTGAGCTGGTGGATCTATGCTCTATCACAGAGATATTTTGACGGGATTTCAAGATG 2199
Qy 2363 CACTGCGGGGTGGCGCTCCGTFCTACTGATCCGAAGCGGCGGCTTCAAGGGTGG 2422
Db 2200 CATTGCGGTGGATGGAGGTCTATTACTGATCCCTAAGAGGCGCTCAATTTCAAGGTTCA 2259
Qy 2423 GCGCCGATCAATCTATCGACCGTCTCAACGAGTGTCTCGGTGGCGCTGGGGTCCGTC 2482
Db 2260 GCTCTTATTAATCTATCAGACAGTTAAACAGGTTTGGGTGGGACTTGGATCGGT 2319
Qy 2483 GAGATCTTCTACGCGGACAGCCCGCTCTCTGACGGCTACGAAGAACGCAACCTCAAG 2542
Db 2320 GAGATATTTTACGCGGACAGTCTCTCTGATGGCTACGAAGGAGCAACTCAAG 2379
Qy 2543 TGGCTGAGCGCTTCCGCTACATCAACACACCATCTACCCCTTCACTCGCTCCCGCTG 2602
Db 2380 TGGCTTGAAGGTTTGGTATGCAACACAAATCTACCCCTTCACTATACCACTT 2439
Qy 2603 CTGCGCTACTGACACCTCCCGCGTCTGCTCTCACCAGGAGTTCAATGCGGTG 2662
Db 2440 CTGCGCTACTGATCTCTCCAGCGATCTGCTCTCTACTGACAAATTCATATGCGCACCG 2499
Qy 2663 ATTAGCAGCTTCCGCGAGCTCTCTTCTCATCGCCCTCTTCTATGTCATCTTCGCGAGCGG 2722
Db 2500 ATAAGCACATTTGCTAGTCTCTTCTTCTCATCTCACTGTTTATGTCGATCATTTGAACGGGA 2559
Qy 2723 ATCTGAGAGTCGGTGGAGCGGGTGAGCATCGAGAGTGTGGAGGAACGAGCAGTTC 2782
Db 2560 ATCTTGAATGAGATGGAGCGAGTTAGCATTTGAGGAATGTTGGAGAAACGAGCAATTC 2619
Qy 2783 TGGGTCTATCGCGCGTGTCCGCGCATCTCTTCCGCGTGTGTGAGGCGCTGCTCAAGTTC 2842
Db 2620 TGGGTCTATGGAGAACTCTAGCTCATCTCTTGGCGTGTGCAAGTCTCTTCAAAATC 2679
Qy 2843 CTCGCGGGATGACACCAACTTCACCGTCACTTCAAGCGCCACCGCGACGAGGACGAC 2902
Db 2680 TTAGCAGGCATTCACAAACTTCACCGTCACTCAATCAAGGCAAC-----AGATGATGAT 2733
Qy 2903 GAGTTCGCGAGCTCTACGCTTCAAGTGACACCGTCTCTCATCCGCGCCACACGCTG 2962
Db 2734 GACTTTGGAACTTTACGCAATTTCAATGGACAAACACTGCTGATCCCTCCCAACACTGTC 2793
Qy 2963 CTCATCATTAACGTCATCGCGCTCGTGGCGGCGATCTCCGACGCCATCAACACGGGTAC 3022
Db 2794 TTAATATAAACAATTTGTCGCTGTTGTCAGGCGATCTCAGATGCCATTAACAATGATAT 2853
Qy 3023 CAGTCTCGGGGCGCCCTCTTCGCGCAAGCTCTTCTTCCGCTCTTGGGTCATCGTCCACCTC 3082
Db 2854 CAGTCTTGGGACCTCTATTGTTGAACCTCTTCTTCTTCTTGGGTCATGTTTCATCTC 2913
Qy 3083 TACCGTTCCTAAGGGGCTCATGGGCGCCAGAACGAGACGCGCCACCGTTGTTGTCATC 3142
Db 2914 TACCCATTCCTAAAGGCTCTGATGGGTGAGACAGAAACAGAACACCAACCATTTGTTGAT 2973
Qy 3143 TGTGTCAATCTGCTGGCTCCATCTCTCTCTCTGCTGCTGGTCAAGATCGACCTTTCATC 3202
Db 2974 TGTGTAGTGTATTGGCATCTATCTCTCTTCTTCTTGGTGAAGAAATGATCTCTTTGTG 3033
Qy 3203 GTCAGGACCAAGGGCGCGGACGTCAGGACAGTGTGGCATCAATTTGCTGA 3250
Db 3034 CTCAGACCAAGGACCTGACACTTCCAAGTGTGGCATCAACTGCTGA 3081

AF088917 3081 bp mRNA linear PLN 02-JUL-1999
LOCUS Arabidopsis thaliana cellulose synthase catalytic subunit (IRX3)
DEFINITION mRNA, complete cds.
ACCESSION AF088917
VERSION AF088917.1 GI:4886755
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1 (bases 1 to 3081)
AUTHORS Taylor,N.G., Scheible,W.R., Cutler,S., Somerville,C.R. and
Turner,S.R.
TITLE The irregular xylem3 locus of Arabidopsis encodes a cellulose
synthase required for secondary cell wall synthesis
JOURNAL Plant Cell 11 (5), 769-780 (1999)
MEDLINE 99264300
PUBMED 10330464
REFERENCE 2 (bases 1 to 3081)
AUTHORS Taylor,N.G., Poindexter,P., Scheible,W., Cutler,S., Somerville,C.R.
and Turner,S.R.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1998) Biological Sciences, 3.614 Stopford
Building, University Of Manchester, Oxford Road, Manchester M13
9PT, UK
FEATURES
source
1..3081
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Landsberg erecta"
/db_xref="taxon:3702"
/chromosome="V"
gene
1..3081
/gene="irx3"
/note="cellulose synthase catalytic subunit"
CDS
1..3081
/gene="irx3"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="AAD32031.1"
/db_xref="GI:4886756"
/translation="MEASAGLVAGSHRNRLVVIHNEEPKPLKNDGGFCETCGDOI
GUTVGDLFVACNCEGFPACRCPCEYERREGTQNCQCKTRKRLGSRVGEDEDE
DIDDEYBNIEHQDKHSAEAMLYKMSYGRGDEDDENGRRFPVPIAGHSGEPV
GGYNGEHLKRVHPYPSSEAGSGGWRERDDWKLQHNLGPEPDDPEMGLIDE
AQPLSRKPIASSKINPYRMVIVARLVAVFLRYRLNPHVDALGLMLTSTVCEIW
FAVSWILDQFPKFPPIERTYLDRLSRYEREGEPNMLAPVDVFTVDPKBPPLVT
SNVLISLAMDYPVEKISCTVSDDGASMLTFESLSEAFARKWPFCKKFSIEPRAP
EMVFTLKVDLQDKVHPTFKERRAMKREYEFKVRINAQVAKASKVPLEGWTMQDT
PWGNTKDHGMIQVFLHSGGFDVEGHELPLVYVSRKRPGFQHHKAGAMNALV
RYAGNLTNPFMLNLDCHYVNSKAVREAMCFMDPOIGKVCYVQFPORPDGIDTN
DRYANRNTVFFDINMKLDGIQGPVTGTCVFKQALYGYEPKPKPKPKMISCGGC
PCFGRRRKNKFSKNDMNGDVAALGGAGDEKHELMFEMNFEKTFGSSIFVSTLMBE
GGVPSSSPAVLKEAHIHISGDEKTEWTELGWIYGSITEDIITFGMHKWRHS
IVCMKRPAPKSPAPINLSRDLNOLRWALGSVEIFSRHSLPWYKYGKGLKWLRF
AYANTVYPTSIPLLAYCILPAICLLTDKFTMPPISTFASLFFISLEWSIIVTGILE
LRWSGVISEWRNEQFWIGISAHLEFAVVOGLLXILAGIDTNFTVTSKATDDDFG
ELYAFKWTLLIPPTLIINIVGVAGISDALNNRGVQSWGPLFGKLFPSFWIHLV
PFLGLMGRQNRTPITVIWINSVLLASIFSLMLVRIIDPFVLTKGPDTSKGGINC"

Query Match 44.1%; Score 1516.8; DB 8; Length 3081;
Best Local Similarity 69.8%; Pred. No. 3.8e-190;
Matches 2210; Conservative 0; Mismatches 862; Indels 96; Gaps 8;
Qy 92 ATGAGAGCCAGCGCGGCTGTGGCGGCTCGCACAAACGGAACGAGCTGGTCTGATC 151
Db 1 ATGGAAGCTAGCGCGGCTCTTGTGCGCGGTTCTCATAAACCGTAATGAACCTAGTCTCAT 60
Qy 152 CGGGGCCACAGAGACCCCAAGCCGCTGCGGGCGGCTGAGCGGGCAGGTGTGCGAGATGTC 211

Db 61 CACAACCATGAGAGCCAAAGCCTCTGAAGAACTCTAGATGGACAATTTCTGTGATATGT 120
Qy 212 GGCAGCAGAGGTGGGGCTCAAGGTGGAGCGGCACTCTTGTGCTGCTGCAAGAGTGGCG 271
Db 121 GGAGATCAGATTGGTTTAAACAGTAGAAGAGACCTCTTCTGTAGCTTGCATAGTGTGGT 180
Qy 272 TTCCCGGTGTGGCGGCTCTAGAGTACGAGCGCGGGAGGSCACGAGAACTGGCCC 331
Db 181 TTTCCGGGTGTAGACCTTGTATGATACGAGAGAGAGAGAAACACAAAATGTGCTT 240
Qy 332 CAGTGAAGAGCGGCTCAAGCGCTCAAGGGGAGCGGAGGTGTGCGGGGACGATGAC 391
Db 241 CAGTGTAAAGACTCTTACAGGCTCTAGAGAAAGCCCAAGTGTGAGGNGATGAAGAC 300
Qy 392 GAGAGGACATTCGACGACTGTGAGCAGAGTTCACATTCGACGACGAGAAATCAGCAGAG 451
Db 301 GAAGAAGATATTGATGATATTGATATGAAATTAATATCGAAATGAACAAGA----- 353
Qy 452 CAGCTGAGGGCAACATGCAGAACAGCCAGATCACCGAGGCGATGCTGCACGGCAGGATG 511
Db 354 -----TAAGCATTAAGCATTTCTGCTGAGGCTATGCTTTATGGGAAAATG 396
Qy 512 AGCTTACGGGAGGGGCGCGAGCGGCGAGCGGCAACACCCCGCAGATCCCGCCCATC 571
Db 397 AGCTATGGAAGAGTCTTGAGATG-----ATGAGATGGGAGATTTCCACCTGTT 447
Qy 572 ATCACCGGCTCCCGCTCCGTCGCGGTGAGCGGTGAGTTTCGATTAACAAGGGTATGGC 631
Db 448 ATAGCTGGTGTCA-----TAGTGGAGAAATTTCCAGTTGGAGGAGTTATGGT 495
Qy 632 CAGGCGAGGTCTGCTTCCCTGCACAGGGCATCCATCCGTACCCCTGTGTCTGAGCCA 691
Db 496 AATGGAGAAC-----ATGGGCTTCAAGCGTGTGCACCCCATATCCATCATCTGAAAGCT 549
Qy 692 GGGAGTGCCTCAAGTGGGACGAGAAAGAGTGAAGTGAAGAGAGATGGAGCACTGG 751
Db 550 GGGAGT-----GAGGAGGATGGCGGAGAAAGATGGATGACTGG 588
Qy 752 AAGTCAAGCAGGGGATCTCTGGCGGCGGCGGATCCGAGAGACATGGAGCGGAGCTG 811
Db 589 AAGCTCAGCATGGAATCTTGGCCAGAACACAGATGATGATC-----CTGAGATG 639
Qy 812 GCATGAACAGAGGGGAGGAGCGGCGGCTGTGAGGAGAGTGTGATGCGGTGAGGCAAG 871
Db 640 GGACTGATGACAGAGGACGAGCAGCCTCTCGCGGAAAGTTCCCATTTGCCCTCAAGCAAG 699
Qy 872 GTGAACCGGTACCGGATGGTGTGCTGTGCTGTGCTGTGCTGTGCTTCTTCTCCGG 931
Db 700 ATCAATCCATATCGGATGGTCACTGTTGTAGGCTTGTGATTTCTAGCAGTTTCTCGGG 759
Qy 932 TACCGTATCTGCAACCCGCTCCCGAGCCCATCGGGCTGTGGCTCGTCTCCATCATCTGC 991
Db 760 TATAGGCTTTGAATCCAGTGCATGATGCTCTGGGATTTATGGCTGACCTCTGTGATCTGT 819
Qy 992 GAGATCTGGTTGGCCATCTCTGATCCTCGACAGTTCCTGACAGTTCCTCCAGTGGTTCCCATCGAC 1051
Db 820 GAAATCTGGTTGGTGTCTCTGATTTCTGATTTCTGATCAGTTTCCCAAGTGGTTCCCTATTGAA 879
Qy 1052 CGCGAGACGTACCTCGACCGCTCTCTCAGTACGAGAGGAGGAGGAGCGGAGCGGTGCTG 1111
Db 880 CGTGAGACCTATCTAGATCGGCTTCTCAGGTACGAGAGAGAGTGAACCAATATG 939
Qy 1112 CTGTGCGGTGGACCTCTTCTGTGAGCAGCGTGGACCCGCTCAAGGAGCGCGGCTGGTG 1171
Db 940 CTTGCCCTGTGATGCTTTGTGATGCTGTGATGCTGTGATTTATGGAGGAGCTCCCTCGTC 999
Qy 1172 ACCGCCAACACGTTGCTCTCATCTCTCGCGGTAGACTACCCGTTGAGCAAGGTCTCTGTC 1231
Db 1000 ACATCCAAACATCTGTGCTGTCAATCTTGGCCATGACTACCCAGTTGAGAAAATCTCTCTGC 1059
Qy 1232 TAGTCTCCGAGCAGCGGCTGATGCTGACGTTGCGAGTTCGAGTTCGAGAGCGGCGGAG 1291
Db 1060 TATGTCTCTGACGAGCGGTGCTTCAATGCTTACATTCGAATCTCTCTCGGAAACTGCTGAG 1119

Qy 1292 TTCCGGCGCAAGTGGGTGCCCTTCTGCAAGAAGTTTCGGCATCGAGCCCCCGGCCCGGAG 1351
Db 1120 TTTTCAAGAAAAATGGGTTCCTTCTGTAAGAAAAATTCCTCATAGAGCCACCGGCCACCGGAG 1179
Qy 1352 TTCTTACTTCTCGCTCAAGGTTCGACTCCTCAAGGACAAAGTGTGACGCCACCTTCGTGCGAG 1411
Db 1180 ATGTACTTTCAGTTGAAAGTTGATTTATCTCCAGGACAAAGTTCACCCCAACAAATTTGTTAAG 1239
Qy 1412 GAGCGCGCCCATGAAGAGAGAGATGAGGAGTTTCAAGGTTCGGATCCGATCAACCGCTGTGTG 1471
Db 1240 GAAGCTCGAGCCATGAAGAGAGATATGAGGAGTTTCAAGGTTCAGGTTCAGATCAACGCTCAAGTG 1299
Qy 1472 GCAAGGCCATGAAGGTTCGCGCAGAGAGGTTGATTCATGAAGGACGGGACGCGCTGGCCCC 1531
Db 1300 GCGAAGCCCTCAAGAGTTTCTCTAGAAAGTTGGATCATGCAAGATGGAACAACCGTGGCCA 1359
Qy 1532 GGGAAACAACCCCGCAGACCCCGGCATGATCCAGGTGTTCTTGGGCCACAGCGGGCGC 1591
Db 1360 GGGAAACAACACCAAGGACCCCGGTATGATCCAAAGTCTTCTCGGCGCACAGCGGAGGA 1419
Qy 1592 CAGACACCGAGGGCAACGAGCTGCCCGCTCGTGTACGTCTCCGTGAGAAGCGCCCG 1651
Db 1420 TTTGATGTCGAGGGCATGAGCTTCTCGGCTTGTGTACGTGTCCCGTGAAGAGGTCTCT 1479
Qy 1652 GATTTCCAGCAACCAAGAAAGCGCGGCCATGAACGCTCTGTATTCGGGTCTCCCGCGTG 1711
Db 1480 GGTTTTCAACACCAACAAAGAGTGGCGCCATGAATGCCCTGGTTTCAGGTGGCAGCGTA 1539
Qy 1712 CTGACCAACCGGCATTCATGCTCAACTTGGACTGTGATCACTACATCAACAACAGCAAG 1771
Db 1540 CTCACAAATGCTCTTTCATGCTGAACTTGGACTGTGATCACTATGATAACAACAGCAAG 1599
Qy 1772 GCCATCCGGAGGCGCATGTGCTTCTCATGAGACCTCAGGTCCGCGCGGAGGCTCTGTAC 1831
Db 1600 GCCGTGAGGGAGCAATGTGTTTTCATGATGATCTCTCAGATTCGAAAGAGTCTGTCTAT 1659
Qy 1832 GTTCAGTTCCCGCAGAGGTTTCGACGGCATCGA CGTGCA GAGACCGATACGCTAA CAGGNAC 1891
Db 1660 GTTCAGTTCCCTCAAGGTTTGTATGGCAATTCACACAAACGATCTGTTACGCCAACAGAAAC 1719
Qy 1892 ACCGTCTTCTTCGACATCAACATCAAGGGCTTGACGGCATCCAAAGCCCGGTGTGAGTC 1951
Db 1720 ACAGTCTTCTTGTATATCAATATGAAAGTCTAGATGGAATCCAAAGTTCAGGTTTACGTT 1779
Qy 1952 GGGACAGGGTGTGTTTCCGGCGCCAGCGCTCTACCGGCTACAAACCTCCCAAGGAGACC 2011
Db 1780 GGTACTGTTGTGTTTCAACCGACAGCTCTGTATGTTATGAACCAACCAAGGGTCTCT 1839
Qy 2012 AAGAGGCCCAAGATGGTGAACCTGCACTGTGCGCGGTGCTTTCGGCCCGCAAGAGCGGAAA 2071
Db 1840 AAACGTCCAAAGATGATAAGCTGTGTTGTTGCTTGTGCTTGGCGCGCGGAGAAAGAAAT 1899
Qy 2072 CA-----CGCAGAGGAGGGCTGCCCGAGGGCACCCCTGATATATGGGATGATAGC 2122
Db 1900 AAGAAATTTTCCAAAGATGACATGAATGGTGAAGTAGAGCCCTTGGAGGAGCAGAGGT 1959
Qy 2123 GACAAGGAGATGCTCATGTCCCATGAACTTCAGAGAAAGGGTTTCGGCGAGTCCGCGCGC 2182
Db 1960 GATAAGGAGCATTTGATGTTTGAATGAACTTTGAGHAAACATTTGGCAATCATCCATC 2019
Qy 2183 TTGCTCAAGTGCAGCTGTGATGGAGGAAGCGGCGTCTCTCTGTCGAGCGCCCGCGCG 2242
Db 2020 TTTGTAACTCAACTTGTATGGAAGAGTGGTGTCTCTCGTCTCATCAAGTCTCTGCAAGT 2079
Qy 2243 CTCTCAAGAGGCCATCCATGTTCATCAGCTGGGCTACGAGGACAGACCGATGGGGG 2302
Db 2080 CTCTTAAAGAGGCAATCCCATGTATGAAGTGGCGTTATGAAAGCAAGACTGAGTGGGGA 2139
Qy 2303 CTGAGCTGGGTGGATCTACGGGTTCATCGAGGACATCTCTGAAGGGGTTCAAGATG 2362
Db 2140 ACTGAGCTGGGTGGATCTATGGCTCTATACAGAGGATATTTTGAACGGGATTCAGATG 2199


```
Db 1830 GTCTGGTGTCTGCTGCTGCTGCTTTCGGCGGGGGAAGCGCGGCAAGTTCGCACAAGAA 1889
QY 2068 ----- 2067
Db 1890 CAAGAGGGCGCGCGCGCGGAGGGCGCGGCTTCGACGAGCGCGCGCGGGCTGCT 1949
QY 2068 -----GAAACAGCCAAAGAGCGGGTTCGAGAGGCAACCGCTGAT 2107
Db 1950 CGGGTTCTACAGAGAGGAGCAAGAGGACAAAGCTCGCGCGCGCGCGCTCGCTGCG 2009
QY 2108 ATGGGAGTAGATAGCAGCAGGA----- 2130
Db 2010 CGGAGGGAAGAAAGGGTACCGGAAGCAGCAGCGGGTTCGAGCTGGAGAGATTCGAGGA 2069
QY 2131 -----GATGCTCATGTCTCCACATGAATTCGA 2157
Db 2070 GGGCTTCGAGGGGTACGACGAGCTGGAGCGCTCGTCTCATGTTCGAGAGAGCTTCGA 2129
QY 2158 GAAGCGGTTTCGGGAGCTTCGCGCGGTTCGTCAAGTTCGACGCTGATGGAGGAAGCGCGGT 2217
Db 2130 GAAGCGGTTTCGGGAGCTTCGCGCGGTTCGTCAAGTTCGACGCTTCACCGCTTCGAGGAGCGCGCT 2189
QY 2218 CCCTC-----CTTCGTCGAGCGCGCGCGGCTTCGAGGAGCGCATCATGTCTCAG 2271
Db 2190 CCCCCAGGGCGCGCGCGGCTTCGAGGAGCGCGCTTCATCAAGGAGGCGCATTCACGTCTCAG 2249
QY 2272 CTGCGGCTACGAGGACAGACCGACTGGGGGCTGGAGCTGGGTGGATCTACGGGTTCGAT 2331
Db 2250 CTGCGGCTACGAGGAGAGACCGAGTGGGCAAGGAGATTGGGTGATCTACGGGTTCGAT 2309
QY 2332 CACGAGGACATCTGAAGGGTTCAAGATGCAATCGCGGGTGGCGCTCCGTTGACTG 2391
Db 2310 GACGAGGACATCTTAACGGGGTTCAAGATGCAATTCGCGGTGGAGTTCGTTGACTG 2369
QY 2392 CATGCCGAGCGCGCGGCTTCAGGGTTCGGGCGCGCATCAATCTATCGACCGCTCTCAA 2451
Db 2370 CACGCGGAGGCGCGGCTTCAGGGTTCGGGCGCGCATCAATCTATCGACCGCTCTCAA 2429
QY 2452 CACGAGTGTCTCGGTTGGCGCTTCGCGGTTCGAGATCTTCGAGCGGCGCACAGCGCCCT 2511
Db 2430 CACGAGTGTCTCGGTTGGCGCTTCGCGGTTCGAGATCTTCGAGCGGCGCATTCGCGCT 2489
QY 2512 CAGTGTACGGGTACAGAGAGCGGCACTCAAGTGGCTGGAGCGCTTCGCTTACATCAACAC 2571
Db 2490 CTGGTACGCTATGG---CGGCGCGCTCAAGTGGCTTCGAGCGCTTCGCTTACACCAACAC 2546
QY 2572 CACCATCTACCCCTTCACCTCGCTTCGCGGTTCGCGGTTCGAGATCTTCGAGCGGCTTCG 2631
Db 2547 CATCGTCTACCCCTTCACCTTCGCTTCGCGGTTCGCGGTTCGAGATCTTCGAGCGGCTTCG 2606
QY 2632 CTTCTCACCGGCAAGTTTCATCATGCGTTCGATTCGAGCTTCGCGGCGCTTCCTTCTTCAT 2691
Db 2607 CTTCTCACCGGCAAGTTTCATCATTCGCGGTTCGAGATCTTCGAGCGGCTTCGAGCTTCAT 2666
QY 2692 CGCCCTCTTCATGTCATCTTCGAGCGGCACTTCGAGATCGGTGGAGCGGGGTGAG 2751
Db 2667 AGCGCTTTCTGTGATCATCGCAGCGGGGTTCGAGCTTCGAGCGGCGGTTCGAGCGGGGTGAG 2726
QY 2752 CATCGAGGAGTGTGGAGGAACGAGCAGTTCTCGGTTCATCGGCGGCGGTTCGCGCATCT 2811
Db 2727 CATCGAGGAGTGTGGAGGAACGAGCAGTTCTCGGTTCATCGGCGGCGGTTCGCGCATCT 2786
QY 2812 CTTTCGCGCTCGTTCGAGGGCTTCGAGGCTTCGCGGCGATCGACACCACTTCACCGT 2871
Db 2787 GTTCGCGGTTCGAGGGCTTCGAGGCTTCGAGGCTTCGCGGCGGTTCGAGCAGCTTCAGGT 2846
QY 2872 CACCTCAAAGGCAACCGGCGCAGGAGCAGAGTTCGCGGCTTCGCGGCTTCAGATG 2931
Db 2847 GAGGTCAAAGGCGCGCGCGCAGGAGCAGGAGCTTCGCGGCTTCGCGGCTTCAGATG 2906
QY 2932 GACGAGCTCTCATTCGCGGCGCAGGAGCTTCGAGGCTTCAGATGAGCTTCGCGGCTTCGCGG 2991
Db 2907 GACGAGCTCTCTGTCGCGGCGCAGGAGCTTCAGGCTTCATCATCAATGATGGTGGGATTCGTCG 2966
```

```
QY 2992 CGGCATCTCGAGCGCATCAACAAACGGGTACCACTCTCGGGGCGCCCTCTTCGCGCAAGCT 3051
Db 2967 CGGCGTGTTCGAGCGCGGTGAACAAACGGGTACGGGTCTGGGGCGCGCTGTCGGGAAGCT 3026
QY 3052 CTTCTTCGCGCTTCGGGTTCATGCTCACTCTCACTCTCACTCTCACTCTCACTCTCACTCT 3111
Db 3027 CTTCTTCGCTCTCTGGGTTCATCTCCACCTCTACCCCTCTCTCTCTCTCTCTCTCTCTCT 3086
QY 3112 CAGAACAGAGAGCGCCACCGCTTCGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3171
Db 3087 GCAGAACCGGACCGCCACCATTCGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3146
QY 3172 CTGCTCTCTGGTTCAGGATCGACCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3231
Db 3147 CCTCGTCTGGTTCAGGATCGACCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3206
QY 3232 GTGTGGCATCAATTCGAGCTGTT 3256
Db 3207 ATCGGGGTCTCTGCTGAGCTGCT 3231

RESULT 12
LOCUS CQ857672 3410 bp DNA linear PAT 31-AUG-2004
DEFINITION Sequence 25 from Patent WO2004070039.
ACCESSION CQ857672
VERSION CQ857672.1 GI:51851797
KEYWORDS
SOURCE
ORGANISM Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Matsfeld, Y. and Broekaert, W.
TITLE Rice promoters
JOURNAL Patent: WO 2004070039-A 25 19-AUG-2004;
Cropdesign N.V. (BE)
FEATURES
source
1..3410
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Query Match 42.8%; Score 1473.6; DB 6; Length 3410;
Best Local Similarity 70.1%; Pred. No. 1.8e-184;
Matches 2290; Conservative 0; Mismatches 674; Indels 302; Gaps 11;

QY 187 GAGCGGCGAGGTGTGCGAGATATGCGCGACGAGGTTCGGGCTCACTGAGCGCGGACCT 246
Db 41 GAAGGAGAGAGCTGTGCGGGGTGTGCGCGAGGAGGTGCGCGGAGGAGACCGGAGGCC 100
QY 247 CTTCTGTCGCTGCAACAGAGTTCGGCTTCGCCGTGTGCGCGGCTTCGCTACGAGTACGAGCG 306
Db 101 GTTCGTCGCGTGTGCGCGAGTTCGGCTTCGCCGTGTGCAAGCCCTGCTACGAGTACGAGCG 160
QY 307 CCGGAGGAGGACCGCAAGACTGCCCGCAGTTCAGAGCGCGCTACAGCGCTTCAGGGGAG 366
Db 161 CAGCGAGGAGGACCGCAGTTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 220
QY 367 CCGGAGGAGTTCGCGGCGGAGCATGACGA---GGAGGAGCATGACGAGCTGAGGAGCACTGAT 423
Db 221 CCACGCGGTGGAAGGCGACGAGGAGCGCGCGCGGACATGAGAGCTTCGAGGAGGAGT 280
QY 424 CAACATCGACGAGCAGAGATTCAGGAGGAGCTTCGAGGAGGAGCAATGCGAGAACAGCCAGAT 483
Db 281 CCAGATCAAGAGAGCGCCACCAAGCAGA-----NACCCCGCCACGAGCGCGCTC 326
QY 484 CACGAGGAGGAGTCTGACGCGGAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 543
Db 327 AACTTCGAGCTCTACTCGGAGAACGCGGAGACGAGCGGAGCAGAGAGTGGCGCTTCGAGG- 385
```


QY 531 --ACGACGGGACGGCAACAACACCCCGAGATCCCGCCCATCATCAACGGCTCCCGCTC 588
Db 727 GTATGACAGTGGCGAGATTCTCTCGGGGATACATCCCATCAGTCACTAAACAGCCAGATCTC 786
QY 589 CGTCCCGGTGAGCGGTGAGTTTCC--GATTACCAACGGGTATGGCCACGGCGAGGTCTCGT 647
Db 787 AGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGCCCAATCGGAAACATTGG 846
QY 648 CTTCCCTGCACAAGCGCATCCATCCGTACCTCTGTCTGAGCAGGGAGTGCCAAAGTGGG 707
Db 847 CAAGCGTCTCCATTTC--CCTATGTGAACCATTCGCCAAATCCCGTCAAGGGAGTTCTCTG 905
QY 708 ACGAGAAAGAGTGAAGTGAAGGAGAGATGGACGACTGGAAGTCCAGACGGGCA 767
Db 906 GTAGCAITGGGAAATGTTCCCTGGAAAGAGAGGGTTGATGGCTGGGAAATGAAGCAGGACA 965
QY 768 -----TCC 770
Db 966 AGGGACGATTTCCCATGACGAATGGCAACAGATTGCTCCCTCTGAGGGTGGGGTGTG 1025
QY 771 TCGCGGGCGGCGCGATCCGGAAGACATGGACCGCGAGCTGGCACTGAACGACGAGGGCA 830
Db 1026 GTGATATTGATGCATCAACTGATTACACATGGAAGATGCCTTATTGAACGACGAACTC 1085
QY 831 GGCAGCCGCTGTGAGGGAAGGTGTCGATCGCGTCGAGCAAGGTGAACCCGTACCGGATGG 890
Db 1086 GACAGCCTCTATCTAGGAAAGTTCACCTCTCTCCAGGATAAATCCATACAGGATGG 1145
QY 891 TGATCGTGGTGGTCTCGTGTGCTCGCTTCTCTCCGTCGATCCGTCATCCGACCCG 950
Db 1146 TCAITGTGCTGCGATTGATTGTTCTAAGACATCTCTTGCACTACCGTATCAAAATCCTG 1205
QY 951 TCCCGGAGCCATCGGGCTGTGGCTCGTCTCATCTCGAGATCTGGTTTCGCCATCT 1010
Db 1206 TGGCGAATGCATACCATTAATGGCTTCTATCTGTTATATGAGATCTGGTTGCTCTT 1265
QY 1011 CCTGATCCTCGACAGTTCCCAAGTGGTTCCTCCATCGACCGGAGAGCTACTCGACC 1070
Db 1266 CGTGGATATTGGATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGGAGAGCTACTTGATA 1325
QY 1071 GCCTCTCCTCAGGTACGAGGGAAGGGAGCGGTCTGCTGTCCGGGTGACCTGT 1130
Db 1326 GGCTGGCATTAAGGTATACCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTGACATTT 1385
QY 1131 TCGTGACGACGGTGCACCGCTCAAGAGACCGCGCTGGTCAACCGCCAAACACCGTCTCT 1190
Db 1386 TCGTCAGTACAGTCGACCCCAATGAAGAGCCTCTCTTGTCTACTGCCAATACCGTGTAT 1445
QY 1191 CCATCTCCCGTACACTACCCCGTGGACAAGGTCTCTGTCATCGTCTTCGACGACGGCG 1250
Db 1446 CCAATCTGCTGTGGATTACCCCTGTGGATAAGGTCTCTTGCTATGATCTGATGGAG 1505
QY 1251 CGTCGATGCTACGTTCCAGTCTGCTGTCGGAGACGGCCGAGTTCGGCGGCAAGTGGTGC 1310
Db 1506 CTGCGATGCTGACATTTGATGCACTAGCTGAGCTTCAGAGTTTGTCTAGGAAATGGGTAC 1565
QY 1311 CCTTCTGCAAGAAGTTCCGCAATCGAGCCCGCGCCCGGAGTCTTACTTCTCGCTCAAGG 1370
Db 1566 CATTTGTTAAGATACAACTTGAACCTAGAGTCTCTGANTGTTACTTCTCCAGAAA 1625
QY 1371 TCGACTACCTCAAGGACAAAGTGCAGCCCACTTGTGTCAGGAGCGCCGCGCCATGAAGA 1430
Db 1626 TTGATTACTTGAAGACAAAGTGACCCCTTCATTGTTTAAAGACCGCCGGGCCATGAAGA 1685
QY 1431 GAGAGTATGAGAGTTCAAGTTCGGATCAACGGCTGCTGGCCAGGCAAGGCAATGAAGTGC 1490
Db 1686 GAGAAATGAAGAAATTCAAAGTTAGGGTAAATGGCTTTGTTGCTTAAGGACACAGAAGTTC 1745
QY 1491 CGGACAGGGGTGATCATGAAGACGGCAGCGCTGCCCGGGAAACAAACCCCGAC 1550
Db 1746 CTGAGGAAGGATGATCATGCAAGATGGCACACCATGCGCAGGAAACAAATACCGGGACC 1805
QY 1551 ACCCGCGCATGATCAAGGTGTTCTTGGGCCACAGCGGGCGGCCACGACACCGAGGCAACG 1610
Db 1806 ATCTCGAAATGATTCAGGTTTTCTTGTGTCAAGTGTGGCTTTGATATCTGAGGCAATG 1865
QY 1611 AGCTGCCCGCCTCGTGTAGCTCTCCCGTGAAGAGCGCCCGGATTCAGACACCAACA 1670
Db 1866 AGCTACCCCGTTTGGTCTATGTTCTTGTGAAAAGCGTCTTGGATTTCCAGCATCAAGA 1925
QY 1671 AGGCGGCGCCCATGAACGCTCTGATTCGCGTCTCCGCGTCTGACCAACCGCGCATTC 1730
Db 1926 AAGCTGGTGCCATGAATGCTCTTGTGTCTCAGCTGTCTTACCAATGACAATACA 1985
QY 1731 TGCTCAACTTGGACTGTGATCACTACATCAACACAGCAAGCCCATCCGGAGGCAATG 1790
Db 1986 TGTGAAATCTTGAATGTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATG 2045
QY 1791 GTTTCCTCATGAGACCTTCAGGTCCGCGGAGGTCGTGCTAGTTCAGTTCCTCCGACAGGT 1850
Db 2046 GTTTCCTTATGGAACCTTAACCTAGAAAGAGTGTCTGCTACGTCTCAGTTTCCCGAGAT 2105
QY 1851 TCGACGGCATCGACGTGCACGACCGATACGCTAAACAGGAAACACCGTCTTCTTCGACATCA 1910
Db 2106 TCGATGCAATTCAGCAGGAATGATCGATATGCCAACAGGAAACACCGTGTTCGATATTA 2165
QY 1911 ACATGAAGGGCTGGACGGCATCAAGCCCGGTGTAGTTCGGACAGGGTGCCTGTTCC 1970
Db 2166 ACTTGAGAGGTCTTGAATGCGCATCCAGGACCAAGTTTATGTGCGAACTGGCTGTGTTTCA 2225
QY 1971 GGGCCAGGGCTCTACGGCTACAACTCCCAAGGACCAAGA----- 2015
Db 2226 ACCGAAACAGCTCTATATGTTTATGAGCCCCCAATTAAGCAGAAGAGGTGGTTCTTGT 2285
QY 2016 -----GGCCCCAAGATGTGTGACCTGCGACTGTGCTGCCCGCCCAAGA 2063
Db 2286 CATCACTATGTGGCGGTAGGAAGGACGAAGCAATCAAGAGAGGCTCGACACAAGA 2345
QY 2064 AGCGGAAACAGCCCAAGAC-----GGGCTGCCGGAGG 2096
Db 2346 AGTCGAGAAGCATGTGCACAGTTCTGTGCCAGTATTCAACCTTTGAAGATATAGAGGAG 2405
QY 2097 GCACCGCTGATATCGGAGTAGATAGCAGACAGGAGATGCTCATGTCCACATGAACTTCG 2156
Db 2406 GAGTTGAAGGGCTGGATTTGAACGACGAGAAATCACTTCTATGTCTCAATAGACCTGG 2465
QY 2157 AGAAGCGGTTTCGGGACGTCCCGCGCTTCGTGACGTGACGCTGATGGAGGAAGCGCGG 2216
Db 2466 AGAAGAGATTGGCCAGTCCGCGAGTTTGTGTGCTTCCACTCTGATGGAGTATGGTGTG 2525
QY 2217 TCCCTCCTGTCGAGCCCGCGCTCTCAAGGAGGCCATCCATGTTCATCACTGCTGCG 2276
Db 2526 TTCTCAGTCCGCAACTCCGAGTCTCTTCTGAAAAGAGCTATCCATGTTATAAGCTGTG 2585
QY 2277 GCTACGAGGACAGACCGGCTGGGGGTGGAGCTGGGGTGGATCTACGGGTGATCAACG 2336
Db 2586 GCTATGAGGACAAAGCTGAATGGGAACTGAGATCCGGGTGGATCTACGGTTCTGTGACAG 2645
QY 2337 AGGACATCTGACCGGGTTCAAGATGCACTGCGCGGGTGGCGTCTCGTGTACTGCAATGC 2396
Db 2646 AAGACATCTCACCGGATTCAGATGCAACGCGGAGGCTGGCGTCTGATCTACTGCAATGC 2705
QY 2397 CGAAGCGGGCGGCTTCAAGGGTTCGGCGCGATCAATCTATCGGACCGCTCAACACGAG 2456
Db 2706 CCAAGCGCCAGCTTCAAGGGGTCTGCCCCCATCAATCTTTTCGAGCCGTCTGAACACGAG 2765
QY 2457 TGCTCCCGTGGCGCTCGGTTCGAGATCTTCTCAGCGGACAGCCCGCTGCTGT 2516
Db 2766 TGCTCCCGTGGCGCTTGGGTTCGTTGAGATCTCTTTCAGCCGGACCTGCCCTCTGTGT 2825
QY 2517 ACGGCTCAAGAAACGGCAACCTCAAGTGGGTGGAGCGCTTCGCCCTACATCAACACCAACA 2576
Db 2826 ACGGCTAC--GGAGGGCGGCTCAAGTTCTTGGAGAGATTCGCGTATCAACACCAACA 2882
QY 2577 TCTACCCCTTCAACCTCGCTCCCGTGTCTGCGCTTACTGCAACCTTCCCGCGCTGCTGCTCC 2636

```

2883 TCTACCGCTCAGCTCCATCCGCTTCTCATCTACTGCTCCTGCGCCGCTCCTGCTGC 2942
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2637 TCACCGCAAGTTTCATATCGCGTCGATTAGCAGTTTCGCGAGCTCTTCTTCATCGCCC 2696
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2943 TCACCGCAAGTTTCATATCCAGAGATCAGCACTTCGCGAGCATCTGTTTCATCTCCC 3002
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2697 TCTTCATGTCATCTTCGCAACGCGCATCTCGGAGATCGCGTGAGCGGGTGAGCATCG 2756
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3003 TCTTCATCTCGATCTTCGCCACGGCATCTCTGGAGATGAGTGAGCGGGTGCGCATCG 3062
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2757 AGGAGTGTGGAGGAACGAGCAGTTCTGGGTTCATCGCGGGGTGTCGCGCATCTCTTCG 2816
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3063 ACAGAGTGTGGAGGAACGAGCAGTTCTGGGTGATCGGGGGCATCTCCGCGCATCTCTTCG 3122
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2817 CCGTCGTGTCAGGGCTCTCTCAAGTCTCTCGCGGATCGACCACTTCACCGTCACTT 2876
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3123 CCGTGTTCAGGGCTCTCTCAAGTCTCTCGCGGATCGACCACTTCACCGTCACTT 3182
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2877 CCAAGGCCCTCG--GACGAGGACGGGACTTCGCGGAGCTGTACATGTTCAAGTGGACGA 3239
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2937 CGTCTCTCATCCCGCCACACAGCTGCTCATCATTAAGTCATCGCGGTGTCGCGCGCA 2996
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3240 CGTCTCTGATCCCGCCACACCATCTCTGATCATCAACCTGGTGGGCGTCTGCGCGCA 3299
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2997 TCTCCGACGGCATCAACAAAGGTTACAGTCTCTGGGGGCGCTCTTCGCGCAAGCTCTTCT 3056
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3300 TCTCTACGCAATCAACAGCGGATACAGTCTGGGGGCGCTCTTCGCGCAAGCTCTTCT 3359
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3057 TCGCTCTTCTGGGTTCATCGTCCATCTCAACCTCTCAAGGGGCTCATCGGGGCGCCAGA 3116
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3360 TCGCTCTTCTGGGTTCATCGTCCATCTCAACCTCTCAAGGGGCTCATCGGGGCGCCAGA 3419
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3117 ACAGGACGCCACCGTTGTTGTCATCTGTTGTCATCTGTCGCTTCATCTTCTCCCTGC 3176
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3420 ACCGACCCCGACCATCGTCTGTCGTCGGGCTCTCTGTCGGGCTCCATCTTCTCTCTGC 3479
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3177 TCTGGGTTCAGATCGACCTCTTCATCTCAGGACCAAGGCGCGGACGTCAGGAGTGTG 3236
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3480 TGTGGGTTCGATCGACCTCTTACACCGCGCTCACTGCGCGGATCCACGACGTGTG 3539
Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3237 GCATCAATGCT 3248
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3540 GCATCAATGCT 3551
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 14
AY639654
LOCUS
DEFINITION
AY639654
Pinus radiata 3911 bp mRNA linear PLN 29-JUN-2004
complete cds.
AY639654
Pinus radiata
AY639654.1 GI:49182339
Pinus radiata (Monterey pine)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE
1. (bases 1 to 3911)
Klein,A.S., Tibbitts,J., Steven,R. and Anthony,B.
Cellulose synthase genes in Conifers: what we know and what we need
to learn
JOURNAL
(in) Hayashi, T. (Ed.);
PLANT CELL WALLS;
(2005) In press
REFERENCE
2. (bases 1 to 3911)
Klein,A.S., Tibbitts,J., Steven,R. and Anthony,B.
Direct Submission
JOURNAL
Submitted (27-MAY-2004) Biochemistry and Molecular Biology,
University of New Hampshire, 46 College Road, Durham, NH 03824, USA
FEATURES
Location/Qualifiers
```

```

source
1..3911
/organism="Pinus radiata"
/mol_type="mRNA"
/db_xref="taxon:3347"
1..3911
/genes="CesA1"
/notes="PrCesA1"
1..168
/genes="CesA1"
169..3423
/genes="CesA1"
/EC_number="2.4.1.12"
/notes="glycosyltransferase"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/db_xref="GI:49182340"
/db_xref="AAT57672.1"
/translacion="MEASGLVAGSHNRNEFVVIHGEHPKPLNTLSGHVCQICGEDV
GLNTDGEFLVACNECGFPVCRPCIEYERMSYGRSCPCNTRYKRGSPRVEDGDDDE
DVIDI EBFENVAQQRNQIIT EAMHLGRMSYGRSCPCNTRYKRGSPRVEDGDDDE
GHSVSGEITPTSYYADNQLANPAMLRVHPSPSEPGSGRIIMDPNMDI GSGYGNVSW
KERGVGSKENKSGOLDMTEGRYOYNGGFAPNEPEDIIDPDMPTDEARQPSLRKVP
IPSSKINPYRMVIVIRLIVLIGIFLRYLLNPKAYGLMWTATSVICWPAISWILQF
PKWLPISRITYLDRSLRYERGEPSMLAPVDLFSVTDPLKEPPLVANTANTLISLV
DYPVDNVCYSDDGASMLTFESLSETSFARKWVFCFKFDIEPRAPETIYFSQKIDY
LKDQFQTFVKERRAMKREYEBFKVINRLVAKASVKPKEGTMQDGTWPWGNNTDH
PGMIOVFLHSGGLDTEGNELPRLVYVSRKRPQFOHKKAGAMNALVRSAVLTNAP
FMNLCDHYINNSKAI REGMCFMMDPOYGRKVCYVQFPQRPEDGIDRNDRYANRNTVF
FDLNMGLDGIQSPVYVGTGCMFRROALYGYGPKPKPKMVTCDCLPCCGPKRKP
KNSSKKSAGIPAPAYNLGIBEGVEGYDDERALLMSQDLEKKFGQSSAFVOSTLME
NGVQPTANPAELKEAHIHVI SCGYEDKTEWGLKGMVIGSVTEDI LTGFMKTRGMR
SIYCMKRAAFKGSAPINLSDRNLQVRLWALASVEIFMSRHCPITWYGGGLKWLRF
AVINTIYPTSLPLIAYCTLPVLSLLTKGFVLPQISFASLFIATLFIATLFI
MWSGVSI EEMWNEQFWI GVSASHFAVIOGLLAKVLAGIDNTFTVTAKSDDGRG
ELYAFKWTLLIPPTLLINLVGVVGVADALNNGFQSWGPLLGLKLPFAPFWIVHLY
PELGLMGNGRNTPTIIVISILLASVFSLSFWIRIDPFLSKVAGPDTKQCGINC"
3424..3879
/genes="CesA1"
3' UTR
ORIGIN
Query Match 39.8%; Score 1371.8; DB 8; Length 3911;
Best Local Similarity 66.8%; Pred. No. 4.5e-171;
Matches 2200; Conservative 0; Mismatches 937; Indels 166; Gaps 10;
Qy 90 CGATGGAGGCCAGCGCGGCTGCTGGCGCGCTCGCACAAACCGGAAACGAGCTGCTGCTGA 149
Db 167 CAATGGAGGCCAGCGCGGCTGCTGGCGCGCTTCTATAACAGAACGAGTCTGCTGCTCA 226
Qy 150 TCCGGGGCCACAGGAGACCCCAAGCCGCTGCGGGCGCTGAGCGGGCGAGGTGTGCGAGATAT 209
Db 227 TCCATGGACATGAGGGGCCGAGCCCTTTGAACACGTTGAGTGGCCACGCTCTGCCAGATTT 286
Qy 210 GCGGCGACGAGTTCGGGCTCACGGTCGACGGCGACCTCTTCGTCGCTGCAACGAGTGGCG 269
Db 287 GTGGCGAGGACGTCGGGCTTAACACAGACGGCGAGCTGTTCTGTTGCTGTTAATGAGTGGC 346
Qy 270 GCTTCCCGGTGTCGGCGCTCTGCTACAGTACAGCGCCGGAGGCGCACGACAGACTGCC 329
Db 347 GGTTCCTGCTGTCGGCGCTGCTATGATGATACAGAGACGAGAGAGAAATCATGCTGCTGCC 406
Qy 330 CCCAGTGCAGAGCGCTACAGCGCTCAAGCGGAGCCGAGGGTTCGCGGGAGCAATG 389
Db 407 CGCAGTGCATATCTCGTTACAGCGTCAAAAAGGAGTCCACGGGTGGAAGGTGACCATG 466
Qy 390 ACGAGGAGCATCGACGACCTGGACGACGAGTTCAACATCGACGAGCAATCAGCAGA 449
Db 467 ATGAAGAAGACGTTGATGATAGAACATGAATTTAATGTGGAGGCTTCAGCAAGAAACA 526
Qy 450 GGCAGCTGGAGGCAACATGCAGAACACGCGAGATCACCGAGGCGATGCTCCACGGCAGGA 509
Db 527 GGCAG-----CAGATCACCGAGGCGATGCTCCACGGACGCA 562
Qy 510 TGAGCTACGGAGGGGCGCCGCGACGAGCGGCAAGGCAACACACCCCGCAGATCCCGCCCA 569
```


Qy	2599	GCTGCTCGCCTA	CTGCAACCTTCCCGCGGCTTGCTCTCCTACCGGCAAGTTCATCATGCC	2658
Db	2778	ACTCATTTGCCCT	ATTATGCACTTCCAGCGCGTCAGTTTCTCAGTGCACAAATTTGTGATCCC	2837
Qy	2659	GTGATTTAGCAGC	GTTCGCCAGCCTTCTTCATCGCCCTCTTCATGTCCTATCTCGGAC	2718
Db	2838	TCAGATCAGTACT	TTTTCGAAAGTCTATTTTATAGCTCTTTTTCATCTCAATTTTGTGCCAC	2897
Qy	2719	GGGCATCCTGGAGAT	GCGGTGGAGCGGTGAGCATCGAGAGTGTGGAGGAACGAGCA	2778
Db	2898	TGGTATTTCTGGAAT	ATGAGTGGAGTGGAGATTTGAAGATGTTGGCGAATGAACA	2957
Qy	2779	GTTCTGGGT	CATCGCGCGGTGTCGGGCACTCTCTCGCCGTGTCAGAGGCGCTGCTCAA	2838
Db	2958	GTTCTGGGT	TATTCGAGGGGTTTCTGCACATTTTTCGAGTTATTCAAGTCTGCTCAA	3017
Qy	2839	GTTCTCGCGGGAT	CGACACCAACTTCACCGTCACCTCCAGGCCACCGCGGACGAGGA	2898
Db	3018	GGTACTGGCAGGC	ATTGATACAAATTTTCACAGTCACGTCACCAAGGCATC-----AGATGA	3071
Qy	2899	CGAGAGTTCGCGAG	CTCTACGGCTTCAAGTGGACCAAGCTCTCTCATCCGCGCCACAC	2958
Db	3072	CGGTGAGTTGGGAA	CTGTATGCAATTAAGACCACTCTCTCATTTCTCTCTCAAC	3131
Qy	2959	GCTGCTCAT	CATTAACTCATCGCGCTCGTGGCGGCATCTCCGACGCCATCAACACGG	3018
Db	3132	CCTGCTTGT	CATCAACTTGTGGGGTGTGTTGTTGGGTAGCAGATCAATCAACATGG	3191
Qy	3019	GTACCACT	CTGGGGGCGCTCTTCGGGCAAGCTCTCTTCCGCTTCTGGGTGTCATGCTCA	3078
Db	3192	ATTTCAGT	CATGGGCTCTCTCTTGGGTAAGCTTTCTTTTGCAATCTGGGTCAATGTGCA	3251
Qy	3079	CCTCTACCGCT	TCTCAAGGGCTCATGGGCGCCAGACAGGACGCCACCGTGTGTGT	3138
Db	3252	CCTGTAT	CTCTTCAAGGCTCATGGGAGCAGAACCCACCATCGTGTGT	3311
Qy	3139	CATCTGTTCCAT	TCCTGCTGCGCTCCATCTTCTCTCCTGCTCTGGGTGAGGATCGACCCCTTT	3198
Db	3312	TATTTGGT	CAATTCCTGCTGGCATCTGTTTCTCTCTTTCTGGTAAGAAATGATCCCTTT	3371
Qy	3199	CATCGTCAGAC	ACCAAGGCGCGGAGCTCAGCAGTGGGATGCGCATCAATTCGTGAGCTGTTTA	3258
Db	3372	CTTGAGTAAAG	TTTAAAGGCCAGATCTAAACAATGTGGCATCACTGCTGATTTCTTTG	3431
Qy	3259	TTA	3261	
Db	3432	ATA	3434	
RESULT 15				
BD236020				
LOCUS		3951 bp	DNA	linear
DEFINITION				PAT 17-JUL-2003
				Materials and method for modification of plant cell wall
				polysaccharides.
ACCESSION				BD236020
VERSION				BD236020.1
KEYWORDS				GI:33045790
SOURCE				JP 2002527056-A/42.
ORGANISM				Pinus radiata (Monterey pine)
				Pinus radiata
REFERENCE				Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS				1 (bases 1 to 3851)
TITLE				Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
JOURNAL				Blockberg, L.N.
				Materials and method for modification of plant cell wall
				Patent: JP 2002527056-A 42 27-AUG-2002;
				GENESIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE
				FORESTS LTD
COMMENT				OS Pinus radiata (radiata pine)
				PN JP 2002527056-A/42
				PD 27-AUG-2002
				PF 08-OCT-1999 JP 2000575985
				PR 13-OCT-1998 US 09/170862, 11-AUG-1999 US 60/148426 PT

Db 905 TGAGCCTGAAGATTATATTGATCCGATATGCAATGACCGATGAAGCAAGGCAGCCACT 964
Qy 841 GTCAGGAAGGTGTCGATCGCGTGCAGCAAGGTGAA CCCGTA CCGGATGATGATCGTGGT 900
Db 965 GTCCGGAAGGTGCAATTCCTTCAAGCAAAATAAATCCATACCGAATGGTCATTGTAAT 1024
Qy 901 CGGTCTGTTGTCGTGCTTCTTCCCTCCGATCGGTATCTGTCACCCCGTCCCGGACGC 960
Db 1025 TCGACTGATAGTGTGGGTATTTTCTCCGCTATCGTCTCTGAAATCCAGTGAAGAATGC 1084
Qy 961 CATCGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGGTTGCGCCATCTCTCGATCCT 1020
Db 1085 ATATGGGCTCTGGGCCATCTCTATCGTTGTGAAATCTGGTTTGCCTTGTGATGATCT 1144
Qy 1021 CGACGATGTCCTCCAAAGTGGTTCCTCCATCGACCGCGAGACGTACCTCGACCGCTCTCCCT 1080
Db 1145 TGATCAGTTCCTCCAAAGTGGTTCCTTATCAGTCTGTAACGATATCTTGATCGACTGTCAAT 1204
Qy 1081 CAGGTACGAGGGAAGGGAGCGCTCGCTGCTGTCGGCGGTGACCTGTTCTGTGAGCAC 1140
Db 1205 AAGGTACGAAAGGAGCGAACCATCAATGCTTGGCACCTGTTGACCTCTTGTGAGTAC 1264
Qy 1141 GGTGGACCGCTCAAGGAGCGCGCTGGTGAACCGCAACACCGTGTCTCTCATCTCGC 1200
Db 1265 TGTAGATCCACTGAAGGAGCTCTCTTGGTTACTTGCCAAATACAGTATATCAATCCTTC 1324
Qy 1201 CGTAGACTACCCGTGGACAAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1325 AGTAGACTACCTGTAGACAATGTCTGTTATGTCTGATGACGAGCGTGTGATGCT 1384
Qy 1261 GAGGTTGAGTGTGTCGGAGCGCGGAGTTCGCGGCGAAGTGGGTGCGCTTCTGCAA 1320
Db 1385 TACTTTTGAATCTCTCTGAGACCTCAGAAATTTGCCAGAAATGGGTGACCAATCTGCAA 1444
Qy 1321 GAAGTTCGGATCGAGCGCGCGCGGAGTCTACTTCTCGCTCAAGGTGCGACTACCT 1380
Db 1445 GAAATTCGACATGAGCCTCGGCTCCGGAATCTTCTCTCAGAAATGACTATCT 1504
Qy 1381 CAAGGACAAGGTGACGCCACTTCTGTCAGAGCGCGCGCCATGAAGAGAGATGA 1440
Db 1505 GAAGGACAAATTTCAACCCACTTTGTCAAGAGCGCGTGCATGAAGAGAGAAATGA 1564
Qy 1441 GGAGTTCAGGTCCGGATCAAGCGCTGGTGGCCAAAGCCATGAAGTGGCGGACGAGG 1500
Db 1565 AGAATTCAGGTGCGCATCAATCGTGGTGGTGAAGGCGCTCTAAAGTGCCCAAGGAAG 1624
Qy 1501 GTGATCATGAAGACGACCGCGTGGCCGGGAA CAACCCCGCACCAACCCCGCAT 1560
Db 1625 ATGACAAATGCAAGACGCTAGCCCTTGGCTGTGTAATATATCCCTGACCAATCTGGTAT 1684
Qy 1561 GATCAGGTGTTCTGGGCCACAGCGCGGCCACGACACCGAGGCGCAACGAGCTGCCCG 1620
Db 1685 GATCAGGTGTTCTGGGTACAGTGGCGGCTCGATACAGAGGCAATGAGCTTCTCG 1744
Qy 1621 CCTGCTGATGTCCTCCGTCGAGAGCGCGCGGATTCAGACACCAAGAGGCGCGCGC 1680
Db 1745 GCTAGTATATGTTCTCTGAGAGAGACCTGGTTTCCAGCATCACAGAGGCGCGGTGC 1804
Qy 1681 CATGAACGCTCTGATTCGCTCTCGCGTGTCTGACCAACCGCGCATTCATGCTCAACTT 1740
Db 1805 CATGAATGCTTGGTTTGGGTTCTGCTGTGCTACCAATGCTCCATTTATGCTGAATCT 1864
Qy 1741 GGACTGTGATCACTACATCAACAA CAGCAAGGCCATCCGGAGGCCATGTGCTTCTCAT 1800
Db 1865 GGATGTTGATCACTACATTAACATAGCAAGGCAATCAGGGAAGGCGATGCTTTATGAT 1924
Qy 1801 GGACCTCAGGTCCGCGGAGGTCTGCTAGTTCAGTTCGCGAGAGGTTTCAGGCGCAT 1860
Db 1925 GGATCCTCAGGTGGGGAAGGTCGTTATGTCCTCAATTCCTCAGAGATTCGATGCTAT 1984
Qy 1861 CGAGCTGCAGACGATACGCTAACAGGAACACCGTCTCTTTCGACATCAACATGAAGG 1920
Db 1985 TGATCGCAATGACCGTTACGCCAATCGAAACACCGGTATTTCTTGATATCAACATGAAGG 2044

Qy 1921 GCTGGACGGCATCCAAAGGCCGGTGTACGTCCGGACAGGGTGTGCTGTTCCGGCCGACGC 1980
Db 2045 TCTGGATGGAATTCAGAGGCTGTATATGTGGGAATCGATGCTGTTCAAGACCAAGC 2104
Qy 1981 GCTCTACGGCTACAAACCTCCCAAGGACCCCAAGAGGCCCAAGATGTGACCTCGACTG 2040
Db 2105 TCTATATGGGTATGGGCTCCCAAGGCCCAAAAGTCCCAAGATGTGACCTGTGATTG 2164
Qy 2041 CTGCCCTGTCTCGGCCGCAAGAG----- 2065
Db 2165 TCTCCCTTGTGGGTCTCTGTAAGAGTCTCCGAAGAAAATAGTAGCAAGAAAAGTGC 2224
Qy 2066 -CGAACAACAGCAAGAGCGGCTGCCGAGAGGCAACCGCTGATATGGAGTA-----GA 2118
Db 2225 AGGAATCCAGCTCCCGCTTACAATCTGGAGCGGATCGAGGAAGAGTAGAAGGTATGA 2284
Qy 2119 TAGCGACAAGAGATGCTCATGTGCCACATGAATCTCGAGAAGCGTTCGSGGAGTCCGC 2178
Db 2285 TGACGAAGAGCATTTGTTGATGAGCCAACTAGACTTCGAGAAGAAATTTGGCCAGTCTTC 2344
Qy 2179 GGGCTTCGTCACTGACGCTGATGGAGGAGCGGCTCCCTCTCTCTGTCGAGCCCGC 2238
Db 2345 AGCTTTGTTCAATCCACTCTGATGGAGATGGTGGTGTTCGCAACAGCAATCCAGC 2404
Qy 2239 CGGCTCTCAAGAGGCGCATCCATGTTCATGTCAGCTGGGCTACGAGGACAAGACCGACTG 2298
Db 2405 TGAATTTTGAAGAGGCTATTCTATGTCATCAGCTGTGGATATGAAGACAAAACGGAATG 2464
Qy 2299 GGGCTGAGCTGGGTTGATCTACGGTCGATCAGGAGGACATCTGAGGGGTTCAA 2358
Db 2465 GGGAAAGAGCTTGGATGGATCTATGGATCAGTCACAGAGACATCTGACTGGAATCAA 2524
Qy 2359 GATGCACTGCGGGGTGGGCTCCGCTGTACTGTCATCCGAAGCGGCGGCTTCAAGGG 2418
Db 2525 GATGCACTGAGGCTGGCGGTCCTTACTGTATGCCCAACAGCAGCATTCAAAGG 2584
Qy 2419 GTCCGCGCCGATCAATCTATCGGACCGTCTCAACAGGTGTCGCGTGGGCGCTGGGGTC 2478
Db 2585 GTCTGCTCCAATCAATCTATCAGACCGTTTGAACACAGGTGTGCGTGGGCTTGGGATC 2644
Qy 2479 CGTCGAGATCTCTTCAAGCGGACACCGCTCTGCTGTAGCGCTTACAAGAACGCAACCT 2538
Db 2645 AGTAGAAATTTTATGAGCAGACATTCGCCCAATCTGATGCTGCTATGGG---GGAGGTCT 2701
Qy 2539 CAAGTGTGTCGAGCGCTTCGCTTACATCAACACCAACCATCTACCCCTTCACTCGCTCCC 2598
Db 2702 GAAATGGCTTGAAGATTTGCTTATCAACACCAATGTCTATCCATTCACCTCTCTTCC 2761
Qy 2599 GCTGCTGCTTACTGCAACCTCCCGCGCTGCTGCTCTCACCGGCAAGTTCATCATGCC 2658
Db 2762 ACTCATTTGCCCTATTCACACTTCAGCGCTCAGTTTGTGCTCCTGCGCAAAATTTGTGATCCC 2821
Qy 2659 GTCGATTAGCACCTTCGCGAGCTCTTCTTCACTGCGCTCTTCACTGCTCATCTTCGCGAC 2718
Db 2822 TCAGATCAGTACTTTTGAAGTCTATTTTAAATAGCTCTTTTCACTCAATTTTTCGCAC 2881
Qy 2719 GGGCATCTCGAGATCGGTGAGCGGGGTGAGCATCGAGAGTGGTGGAGGAACGAGCA 2778
Db 2882 TGGTATCTGAANAATGAGGTGGAGTGAGCATTTGAAGAAATGGTGGCGAATGAACA 2941
Qy 2779 GTTCTGGGTCAATCGCGGCGTGTCCGCGCATCTTCTGCGCTGCTGTCAGGCGCTGCTCAA 2838
Db 2942 GTTCTGGGTATTTGGAGGGGTTTCTGCACATTTTGTGAGTATTTCAGAGTCTGCTCAA 3001
Qy 2839 GGTCTCTCGCGGATCGACACCACTTCAACGCTCACTCCAAAGGCCACCGGAGACGAGGA 2898
Db 3002 GGTACTGGCAGGATGATACAAATTTCAAGTCACTGCGCAAGGCATC-----AGATGA 3055
Qy 2899 CGACGAGTTCGCGAGCTCTACGCTTCAAGTGGACCAACCGCTCTCATCTCCGCGCCACAC 2958
Db 3056 CGGTGAGTTTGGGGAACGTGATGCAATCAATGGACCACTCTCTCATTTCTCTCTACAC 3115

QY	2959	GCTGCTCATATTAAAGTCATCGCGCTCGTGGCCGGCATCTCCGACGCCCATCAACAACGG	3018
Db	3116	CCTGCTTGTTCATCAACCTTGTGGGGTGGTTGTGGCGTAGCAGATGCAATCAACAATGG	3175
QY	3019	GTACCAAGTCCTGGGGGCCCTCTTCGGCAAGCTCTTTCGCCCTTCTGGGTCAATCGTCCA	3078
Db	3176	ATTTCAGTCATGGGGTCTCTCTTGGGTAAAGCTTTTCTTGCATTTCTGGGTCAATTGTGCA	3235
QY	3079	CCTCTACCGTTTCCTCAAGGGGCTCATGGGGCCGAGAACAGGACGCCACCGTCTGTGT	3138
Db	3236	CCTGTATCTTTCCTCAAGGTCTCATGGCAGGAGAACCCGACCATCGTGGT	3295
QY	3139	CATCTGGTCCATTCTCTGGCCCTCCATCTTCTCCCTGCTCTGGGTCAAGATCGACCTTT	3198
Db	3296	TATTTGGTCAATTCTCTGGCATCTGTTTCTCTCTTTCTGGGTAAGAATTGATCCTTT	3355
QY	3199	CATCGTCAGGACCAAGGGCCCGACGTCAGGCAAGTGTGGCATCAATTGCTGAGCTGTTTA	3258
Db	3356	CTTGAGTAAGGTTAAAGGCCCAGATATAACAATGTGGCATCAACTGCTGATTTCTTTG	3415
QY	3259	TTA 3261	
Db	3416	ATA 3418	

Search completed: March 11, 2005, 21:58:40
Job time : 9704 secs

Db 161 CAGGAGGGCACCCAGTGTGCTGCCCCAGTGCACACCCCGCTACAGCGCCACAAAGGGTG 220
Qy 367 CCCAGGGTTGTCGGGGACGATACGA- ---G3AGGACATCGACACCTGAGCAGCGATT 423
Db 221 CCCAGGGTTGAAAGCGACGAGGACGACGCGCGACATGACGACTTCGAGGAGGATT 280
Qy 424 CAACATCGACGACGAGNATCAGCAGGAGGAGCTGAGGGGCAACATGCAAGAACGCGCAT 483
Db 281 CCAGATCAAGAGGCCCAACGAAGCAGA- ----AACCCGCCACGAGCGCGTC 326
Qy 484 CACGAGGGCATGCTGCACGGCAGGATGAGTACGGGAGGGGCCCGACGACCGCGCAGCG 543
Db 327 AACTTCGACGTCTACTCGGAGAACGGCGAGCAGCGGCAAGAAAGTGGCGCCCTGAGG- 385
Qy 544 CAACAAACACCCCGCAGATCCCGGCCATCATCACCGGCTCCCGTCCGGTGCAGCGG 603
Db 386 -----CCGGCGCTCTCTTCTTCAACCGAAGCGTGGTGGAGG- 426
Qy 604 TGAGTTTCCGATTACCAACGGGTATGCGCACGGGAGGTCTCGTCTTCCCTGCACAAGCG 663
Db 427 ----- 426
Qy 664 CATCCATCCGTACCCTGCTGTCTGAGCCAGGGAGTGCCAGTGGGACGAGAAAGNAGT 723
Db 427 -----ATCTGGAGCAGGAGGAGATGGAGGTGGCATGGATGGAGGACAGGATCGA 481
Qy 724 GAGCTGGAAGGAGGATGACGACTGGAAGTCCAAGCAGGGCATCTCCGGCGCGCGCG 783
Db 482 CAAGTGGAG- --ACGAAGCAGGAGAGCGGGCAAGTCAACCGGACGACAGCGACGA 538
Qy 784 CGATCCGAAGACATGAGACGCCGACGTGGCACTGGAACGACGAGGAGGAGCGCGTGTG 843
Db 539 CGACGACGAAGAAGACGACGAGTACATGCTGCTCGCGAGGCGGAGGAGCGCGTGTG 598
Qy 844 GAGGAAGTGTGCATCGCTCGACGAGGTGAACCCGTACCGGATGCGATGCTGTGGTGG 903
Db 599 GAGGAAGTGTCCGATCCCGTCGAGCAAGATCAACCCGTACCGGATGCGATGCTGTCTCG 658
Qy 904 TCTGTTGTGCTCGCCCTTCTTCTCGGTACCGTATCTCTGACCCCGTCCCGGACGCCAT 963
Db 659 GCTGGTGGTCTCTGCTTCTTCTCAAGTTCGGATCAACGCGCGCGATGACGCGGT 718
Qy 964 CGGCGTGTGCTGTCTCCATCATCTCGGAGATCTGGTTCGCCATCTCTTGGATCTCTCGA 1023
Db 719 GCGCGTGTGCTGTGCTGTGATCTGCGAGCTGTGGTTCGCGTGTGCTGTGATCTCTCGA 778
Qy 1024 CCAGTTTCCCAAGTGGTTCCCATCGACCGCGAGAGTACTCTGACCGCTCTCCCTCAG 1083
Db 779 CCAGCTGCCCAAGTGTGCGCGGTGACGAGGAGAGTACTTGACCCGCTGGCCCTCCG 838
Qy 1084 GTACGAGAGGAAGGGAGCGGTGCTGTGCTGTGCGGAGTGGACCTGTTCTGAGCACGGT 1143
Db 839 GTACGAGCGCAGCGGCGAGCGGTGCGCTGCGCCCGATCGATTTCTTCTGACGACGGT 898
Qy 1144 G3ACCCGCTCAAGAGCCGCGTGTGACCGGCAACACCGTGTCTCCATCTCCGCGT 1203
Db 899 G3ACCCGCTCAAGAGCCGCGCATCATACCGGCCAACACCGTGTCTCCATCTCCGCGT 958
Qy 1204 AGACTACCCGTCGACAAAGTCTCTGCTACGCTCTCGAGAGCGGCGCTGATGTGAC 1263
Db 959 CGACTACCCGTCGACCGCGTCTCTGCTACGCTCTCGAGCAGCGGCGGTCATGTGCT 1018
Qy 1264 GTTTCGATGCTGTGCGAGACGCGGAGTTTCGGCGCAAGTGGGTGCCCTTCTGCAAGAA 1323
Db 1019 CTTTCGACACGCTCTCGAGACCGCGAGTTTCGCGCGCGGTGGTCCCCCTTCTGCAAGAA 1078
Qy 1324 GTTCGGATTCGAGCCCGCGCGGAGTTCTATTCTTCGCTCAAGTGCATCTACTCAA 1383
Db 1079 GTTCACCATCAGCCCGCGCGCGAGTTCTACTTCTCCAGAAGATCGACTACTCTCAA 1138
Qy 1384 G3ACAAAGTTCAGCCCACTTCTGTGAGGAGCGCGCGCATCAAGAGAGGTATGAGGA 1443
Db 1139 G3ACAAAGTTCAGCCCACTTCTGTCAAAGAAACCGCGCGCATGAAGAGAGGTATGAGGA 1198

1444 GTTCAAGGTTCCGATCAACCGCTGTGTGCCAAGGCCATGAAGGTTCCCGSCAAGGGGTG 1503
Db 1199 GTTCAAGGTTGAGGATAAACCGCTGTGTGCGAAGCGCGCAGAAAGCCGCGAGGAAGGGTG 1258
Qy 1504 GATCATGAAGGACGCGACCGCTGTGGCCCGGGAACAACACCCCGGCAACACCCCGGCATGAT 1563
Db 1259 GGTGATGCGAGGACGCGACCGCATGGCCGGGGAACAACACGAGGGACACCCCGGGGATGAT 1318
Qy 1564 CAGAGTGTTCCTGGGCCACAGCGCGGCCACACACCCGAGGGCAACAGAGCTGCCCGCCT 1623
Db 1319 CAGAGTGTACTTGGGCGAGCAGGCGCGCTTCGACGTCGAGGGCAGGAGCTGCCGCGCT 1378
Qy 1624 CGTGTACGTTCTCCGTGAGAAGCGCCCGGATTTCCAGCACACCACAAGAGCCCGCGCGCAT 1683
Db 1379 GGTGTACGTTCTCCGCGAGAGCGGCCCGCTACACACCAACAAGAGGCGCGCGCAT 1438
Qy 1684 GAACGCTGTGATTCGGGTCTCCGCGTGTGTGACAAACGCGGCCATTTCATCTCAACTTGA 1743
Db 1439 GAACCTCCCTCGTTCCGCTCTCCGCGTGTCTTACCAACGCGCCCTTCATCTCAACCTCGA 1498
Qy 1744 CTGTGATCTACTATCAACAACAGCAAGGCCATCCGGGAGGCCATGTCTTCCTCATGGA 1803
Db 1499 CTGCGACCTACTAGTCAACAACAGCAAGGCCGCTCCGCGAGGCCATGTCTTCCTCATGGA 1558
Qy 1804 CCTCTAGGTCGCGCGGAAGGCTGTCTACGTTCCCGCAGAGGTTTCGACGCGCATCGA 1863
Db 1559 CAAGCAGCTCGGCAAGAGCTGTGCTACGTTCCAGTTCCCGCAGCTTCGACGCGCATCGA 1618
Qy 1864 CGTGACGACCGATAGCTAAACAGNAACACCGTCTCTTCGACATCAACATGAGGGGCT 1923
Db 1619 CGGCCACGATCGCTACGCCAACCGCAACACCGCTTCTTCGACATCAACATGAGGGGCT 1678
Qy 1924 GGAACGATCCAAAGGCCCGGTGTACGTCGCGGACAGGTCGTGTTCCGCGCGCAGCGCT 1983
Db 1679 GACCGGATACAGGGGCCGCTGTACGTGGGAGCGGGAACGTTTCAACAGGACGCGCT 1738
Qy 1984 CTACGCTACAACTCTCCAAAGGACCCAAAGAGGCCCAAGATGTTGACC- ---- 2032
Db 1739 GTACGGATACGACCGCGCGCGCGGAGAGAGGCCGAAGATGACGTGCGACTGTGCGCC 1798
Qy 2033 -----TCGACTGTCTCCGCTTTCGCGCGCAAGAGCG- ---- 2067
Db 1799 GTCTGTGTCTGTCTGTCTGTCTTCGCGCGGGGAAGCGCGCAAGTCGCAAGAA 1858
Qy 2068 ----- 2067
Db 1859 CAAGAGGCGCGCGCGCGCGAGGCGCGCTCTGACGAGCCGCGCGCGGCTGCT 1918
Qy 2068 -----GAAACACGCCAAGGACGCGCTGCGGAGGCGACCGCTGAT 2107
Db 1919 CGGTTTCTAAGAAGAGGAGCAAGAAGCAAGCTCGCGCGCGCGCGCTGCTGCG 1978
Qy 2108 ATGGGAGTAGATAGCAAGA- ---- 2130
Db 1979 CGGAGGAAGAAAGGTTACCGGAAGCACACGCGCGGTTTCGAGCTGAGAGATCGAGGA 2038
Qy 2131 -----GATGCTCATGTCCACATGAATTCGA 2157
Db 2039 GGGCTCTGAGGGGTACGACGAGCTGGAGCGCTGCTGCTCATGTCCAGAGAGCTTCGA 2098
Qy 2158 GAAGCGGTTTCGGGCACTCGCGCGGTTTCGTCACTGACGCTGATGATGAGGAGCGCGCT 2217
Db 2099 GAAGCGGTTTCGGCACTCGCGCGGTTTCATCGCTCCACCTTCGTCGAGACGCGCGCT 2158
Qy 2218 CCTTC-----CTTGTGAGCCCGCGCTTCCTCAGAGGCCATCCATGTCTATCAG 2271
Db 2159 CCCCCAGGCGCGCGCGCGCGCTTCATCAAGAGGCCATCCACGTCTATCAG 2218
Qy 2272 CTGCGGCTACGAGGAACAAGACCGACTGGGGGCTGAGCTGGGGTGGATCTACGGGTCTGAT 2331
Db 2219 CTGCGGCTACGAGGAGAGACCGAGTGGGGCAAGAGATTGGTGGATCTACGGGTCTGAT 2278


```
Qy 2337 AGGATCTCTGACCGGGTTCAAGATGCACTGCCGCGGTGCGCTCCGTGTACTGCAATGC 2396
Dy 2647 AAGACATTTCTCACCGGATTCAGATGACACGCGGAGGTGCGGTGCGATCTACTGCAATGC 2706
Qy 2397 CGAAGCGGCGGCGTTCAGGGGTGCGGCGGATCAATCTATCGACCGTCTCAACACG 2456
Dy 2707 CCAAGCGGCGGCGTTCAGGGGTGCGGCGGATCAATCTATCGACCGTCTCAACACG 2766
Qy 2457 TGTCTCGGTGCGGCGTCCGCTCGAGATCTTCTTCAGCGGCGACAGCCCGCTGCTGT 2516
Dy 2767 TGTCTCGGTGCGGCGTCCGCTCGAGATCTTCTTCAGCGGCGACAGCCCGCTGCTGT 2826
Qy 2517 ACGGCTACAGAAACGCGAACCTCAAGTGGCTGGAGCGCTTCGCTTACATCAACACCA 2576
Dy 2827 ACGGCTAC--GGAGGGCGGCTCAAGTTCTGGAGAGATTGCGCTACATCAACACCA 2883
Qy 2577 TCTACCCCTTCACCTCGTCCGCTCGCTCGCTACTGACCCCTCCGCGCGTCTGCTCC 2636
Dy 2884 TCTACCCCTTCACCTCGTCCGCTCGCTACTGACCCCTCCGCGCGTCTGCTCC 2943
Qy 2637 TCACCGGCAAGTTTCATCATGTCGCTCGATTAGCAGTTTCGCGAGCTCTTCTTCATCGCC 2696
Dy 2944 TCACCGGCAAGTTTCATCATGTCGCTCGATTAGCAGTTTCGCGAGCTCTTCTTCATCGCC 3003
Qy 2697 TCTTCATGTCATCTTCGCGAGCGGCTCTGCGAGATGCGGTGAGCGGGTGAGCATCG 2756
Dy 3004 TCTTCATGTCATCTTCGCGAGCGGCTCTGCGAGATGCGGTGAGCGGGTGAGCATCG 3063
Qy 2757 AGGAGTGTGAGGAGAACGAGAGTTCTGGGTTCATCGCGCGGCTCTCGCGATCTCTTCG 2816
Dy 3064 ACAGTGTGTGAGGAGAACGAGAGTTCTGGGTTCATCGCGCGGCTCTCGCGATCTCTTCG 3123
Qy 2817 CCGTGTGTCAGGGCGCTGCTCAAGTCTCTGCGCGGATCGACCACTTCACCGTCACT 2876
Dy 3124 CCGTGTTCAGGGCGCTGCTCAAGTCTCTGCGCGGATCGACCACTTCACCGTCACT 3183
Qy 2877 CCAAGGCGACCGGCGACGAGGACGAGTTTCGCGAGCTCTACGCTTCAAGTGGACCA 2936
Dy 3184 CCAAGGCGCTCG--GACGAGGACGCGACTTCGCGGAGCTGTACATGTTCAAGTGGACGA 3240
Qy 2937 CGCTCTCTACCCCGCCACGAGCTGTCTCATTAAGTTCATCGGCGTGTGCGCGCA 2996
Dy 3241 CGCTCTCTACCCCGCCACGAGCTGTCTCATTAAGTTCATCGGCGTGTGCGCGCA 3300
Qy 2997 TCTCCGAGCGCATCAACAAACGGGTACAGTCTCTGCGGCGCTCTTCTGCGCAAGCTCTTCT 3056
Dy 3301 TCTCTACGCGCATCAACGCGGATACAGTCTGTGGGCGCGCTCTTCTGCGCAAGCTCTTCT 3360
Qy 3057 TCGCTCTCTGGGTTCATGTCACCTCTACCTTCTCAAGGGGCTCATGGGCGCGCAGA 3116
Dy 3361 TCGCTCTCTGGGTTCATGTCACCTCTACCTTCTCAAGGGGCTCATGGGCGCGCAGA 3420
Qy 3117 ACAGAGCGCCACGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3176
Dy 3421 ACCGACCGCGACCATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Qy 3177 TCTGGGTTCAGGATCGACCTTTCATCGTCAAGGCGCGGCGCGGCGCTCAGGCGAGTGTG 3236
Dy 3481 TGTGGGTTCAGGATCGACCTTTCATCGTCAAGGCGCGGCGCGGCGCTCAGGCGAGTGTG 3540
Qy 3237 GCATCAATTTGCT 3248
Dy 3541 GCATCAATTTGCT 3552
```

RESULT 4

AAZ99527

ID AAZ99527

XX standard; DNA; 3746 BP.

AC AAZ99527;

XX

DT 03-JUL-2000 (first entry)

XX

DB

XX

KW

XX

OS

XX

FH

FT

FT

FT

FT

FT

FT

XX

PN

XX

PD

XX

PF

XX

PR

XX

PA

PI

PI

XX

DR

DR

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

SQ

DNA encoding a maize cellulose synthase.

Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

transgenic plant; plant breeding marker; ss.

Zea mays.

Location/Qualifiers

321..3449

/*tag= a

/product= "cellulose synthase"

/transl_except= (pos: 1800..1802, aa: Xaa)

/note= "no termination codon given; Xaa is an unspecified

amino acid"

WO200009706-A2.

24-FEB-2000.

16-AUG-1999; 99WO-US018760.

17-AUG-1998; 98US-0096822P.

(PION-) PIONEER HI-BRED INT INC.

Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

WPI: 2000-224343/19.

P-PSDB; AAY84119.

New genes which encode maize cellulose synthase polypeptides in plants

useful for modulating the expression of cellulose synthase in plants and

to produce transgenic plants expressing the novel protein.

Claim 1; Page 176-181; 119pp; English.

The present sequence encodes a maize cellulose synthase polypeptide. The

cellulose synthase can be used for the improvement of stalk quality for

improved stand or silage. It also provides an increased concentration of

cellulose in the pericarp, hardening the kernel and improving its

handling ability. The sequences are used to produce transgenic plants and

seeds expressing the cellulose synthase. The polynucleotide is used for

modulating, preferably increasing, the level of the synthase in a plant

cell. The plants are preferably monocots. The polynucleotide is also used

as a probe or primer in the detection quantitation or isolation of gene

transcripts. The probes are useful in detecting deficiencies in the level

of mRNA in screenings for desired transgenic plant, for detecting or

mutations in the gene, for monitoring upregulation of expression or

changes in enzyme activity in screening assays of compounds, for

detection of any number of allelic variants of the gene, or for use as

molecular markers in plant breeding programs. The isolated nucleic acids

of the present invention can also be used for recombinant expression of

their encoded polypeptides or for use as immunogens in the preparation

and/or screening of antibodies. The proteins can be employed in assays

for enzyme agonists or antagonists of enzyme function or for use of

immunogens or antigens to obtain antibodies specifically immunoreactive

with a protein

Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;

Query Match 40.6%; Score 1398.8; DB 3; Length 3746;

Best Local Similarity 67.2%; Pred. No. 3.8e-222;

Matches 2185; Conservative 1; Mismatches 923; Indels 143; Gaps 9;

Qy 133 GAAGGAGCTGTGCTGATCTCGGGGCGGACGAGGACCGGCTGCGGGCGCTGAGCGG 192

Dy 308 GAGCTCGCTTGTCCATGGAGGCGGACGCGGAGCGGCTGAGCTCGGGGAGCGCGGTGCGG 367

Qy 193 GCAGGTGTGCGAGATATGCGGCGGACGAGGTCTCGGCTCACGCTGACGCGGACCTCTTCGT 252

Dy 368 ACAGGTGTGCGAGATCTGCGGCGGACGCGGTGCGGACGCGGCGGAGGAGCTCTTCGC 427

Qy	253	CGCCTGCACGAGTGGCGGCTTCCCGCTGTGTCGCGGCCCTGTCTACAGGTACGAGCGCCCGGGA	312
Db	428	CGCCTGGACGCTCTGCGGGTTTCGCGTGTGCGGCCCTGTCTACGAGTACGAGCGCAAGGA	487
Qy	313	GGGCACGCAGAACTGCCCCAGTCAAGACGCGCTACAAGCGCTCAAGGGGAGCCCGAG	372
Db	488	CGGCACGAGCGGTGCCCCAGTGCAGACCAGTACAAGCGCCACAAGGGGAGCCCGCGC	547
Qy	373	GGTTGCCGGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAACATCGA	432
Db	548	GATCCGTGGGAGGAGGAGACGACATGATGCCGATAGCGACTTCAATTACCTTGCATC	607
Qy	433	CGACGAGAACTACGACGAGGCGAGCTGGAGGGCAACATGCAGAAACAGCCAGATCA	485
Db	608	TGGCAATCAGGACCAGAACGAGAGAATTGCCGACAGAAATGCGCAGCTGGCGCATGAACGT	667
Qy	486	-----CCGAGGCGATGCTCACGGCAGGATGAGCTACGGGAGGGGCCCCCG-	530
Db	668	TGGGGGAGCGGGGATGTTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAA	727
Qy	531	--ACGACGGGACCGGCAACACCCCGCGAGATCCCGCCCATCATACCGGCTCCCGCTC	588
Db	728	GTATGACAGTGGCGAGATTCTCGGGGATACATCCCATCAGTCACTAACAGCCAGCATCTC	787
Qy	589	CGTGCCGGTGAGCGGTGAGTTTCC-GAATACCAACGGGTATGGCCACGGCGAGGTCTCGT	647
Db	788	AGGAGAAATCCCTGGTGTCTCCCTGACCATATGATGTGCCCACTGCGAACATTTG	847
Qy	648	CTTCCCTGCACAGGGCATCCATCCGTACCTGTGTCTGAGCCAGGGAGGTGCCAGTGGG	707
Db	848	CAAGCGTGCTCCATTTTC-CTATGTGAACCAATTCCGCCAAATCCCGTCAGGGAGTTTCTCTG	906
Qy	708	ACGAGAAAGAAAGTCAAGCTGGAAAGAGAGGATGGAGCAGCTGGAAGTCCAAAGCAGGGCA	767
Db	907	GTACCATTTGGGAATGTTCCCTGGAAAGAGAGGGTTGATGGCTGGAAATGAAGCAGGACA	966
Qy	768	-----TCC	770
Db	967	AGGGACGATTTCCCATGAGAAATGGCAACAGCATTTGCTCCCTCTGAGGGTCGGGTGTTG	1026
Qy	771	TCGCGGGGGCGCATCCGAGACATGAGAGCCGAGGTGGCCTCTGAAACGACGAGGCGGA	830
Db	1027	GTGATATTGATGCAATCAACTGATTAACAATGGAAGATGCTTATTGAAACGACGAAACT	1086
Qy	831	GGCAGCCGCTGTCCAGGAAGGTGTGCGTCCGTCGACAAAGGTGAACCCGTACCGGATGG	890
Db	1087	GACAGCCTCTATCTAGAGAAATTCACCTTCCCTCCTCAGGATTAATCCATACAGGATGG	1146
Qy	891	TGATCGGTGTGCTCTGTTGTGCTCGCCTTTCTCCCGGTACCGGTATCCTGCAACCCCG	950
Db	1147	TCATTGTGCTGCGATTGATTGTTCTAAGCATCTTCTTGGCATACCGTATCACAAATCCTG	1206
Qy	951	TCCCGGAGGCCATCGGGGTGTGGCTGTCTCCATCATCTGGAGATCTGGTTTCGCCATCT	1010
Db	1207	TGCGCAATGCAATACCCATTATGGCTTCTATCTGTTATATGTGAGATCTGGTTTGCCTCTT	1266
Qy	1011	CCTGGAATCTCGACAGTTTCCCAGTGGTTTCCCACATCGACCGCGAGACCTACTCGACCC	1070
Db	1267	CGTGGATATTGGATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGGAGACGTACTTGATA	1326
Qy	1071	GCCTCTCCCTCAGGTACGAGAGGAAGGGAGCGCGTGGCTGTCTCGCGCGGTGGACCTGT	1130
Db	1327	GGCTGGCAATTAAAGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGTCTTTGACATTT	1386
Qy	1131	TCGTGAGACGGTGGACCCGCTCAAGGAGCCCGCGTGGTGACGCGCAACACCGTGTCTCT	1190
Db	1387	TCGTGAGTACAGTCCGACCCCAATGAAGAGGCTCTCTTGTCTACTGCGCAATACCGTGTCTAT	1446
Qy	1191	CCATCTTCGCGTGAATACCCCGTGACAGAGTCTCCTGTAAGTCTCTCGAGCAGCGCG	1250
Db	1447	CCAATCTGTGTGGAATTACCCCTGTGATAAGGTCTCTTGTCTATGTATCTGATGATGGAG	1506
Qy	1251	CGTCGATGTGACGTTTCAAGTCTGCTGTGCGAGACGGCGGAGTTTCCGCGCAAGTGGGTGC	1310


```
Db 2587 GCTATGAGCAAGACTGATGGGGAACGTGAGATCGGTGGATCTACGGTCTGTGACAG 2646
Qy 2337 AGGACATCTGACGGGTTCAAGATGCACTGCCGGGTTGGCTCCGTGACTGATGC 2396
Db 2647 AAGACATTTCTACCGGATTTCAAGATGCACTGCCGGGTTGGCTCCGTGACTGATGC 2706
Qy 2397 CGAAGCGGGCGGCTTCAAGGGTTCGGCGCGATCAATCTATCGGACCGTCTCAACACAG 2456
Db 2707 CAAAGCGGCGAGCTTCAAGGGGTTGCCCCCATCAATCTTTCGAGACCGTCTGAACACAG 2766
Qy 2457 TGCTCCGGTGGGCGTCCGGGTCCTGAGATCTTCTTCAGCGCGGACAGCCCTCTGCTGT 2516
Db 2767 TGCTCCGGTGGGCTCTTGCTGCTCCGTGGAGATCTTCTTCAGCGCGACTGCCCCCTGTGT 2826
Qy 2517 ACGGTACAGAAACCGCAACCTCAAGTGGCTGGAGCGTTCGCTTACATCAACACACCA 2576
Db 2827 ACGGCTAC---GGAGGGGGGCTCAAGTTCCTGGAGAGATTGCGGTACATCAACACACCA 2883
Qy 2577 TCTACCCCTTCACTCGCTCCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2636
Db 2884 TCTACCCGCTCAGTCCATCCGCTCTCTCATCTACTGATCTGCTGCTGCTGCTGCTGCTG 2943
Qy 2637 TCACCGCAAGTTCAATCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2696
Db 2944 TCACCGCAAGTTCAATCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3003
Qy 2697 TCTCATCTCATCTTCGCGACGGGATCTTGAGATCGGTTGGAGCGGGTGGAGCATCG 2756
Db 3004 TCTTCATCTCATCTTCGCGACGGGATCTTGAGATCGGTTGGAGCGGGTGGAGCATCG 3063
Qy 2757 AGGAGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2816
Db 3064 ACGAGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3123
Qy 2817 CCGTCTGTCAGGGGCTGCTCAAGGTCCTCGCGGATCGACCACTTCAACCGTCACT 2876
Db 3124 CCGTCTGTCAGGGGCTGCTCAAGGTCCTCGCGGATCGACCACTTCAACCGTCACT 3183
Qy 2877 CAAAGGCCACGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2936
Db 3184 CAAAGGCCCTG---GACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240
Qy 2937 CGCTCTCATCTCCGCGGACGAGGCTGCTCATCATTAAGTATCGGCGTCTGGCGGCA 2996
Db 3241 CGCTCTCATCTCCGCGGACGAGGCTGCTCATCATTAAGTATCGGCGTCTGGCGGCA 3300
Qy 2997 TCTCGAGCGCATCAACACGGGTACCACTGCTGGGGGCTCTTTCGGCAAGCTTCTTCT 3056
Db 3301 TCTCTACGCCATCAACAGCGGATACCACTGCTGGGGGCTCTTTCGGCAAGCTTCTTCT 3360
Qy 3057 TCGCTCTTGGGTATCTGCTCACTCTTACCGGTTCTCAAGGGGCTCATGGGGGCGGCA 3116
Db 3361 TCGCTCTTGGGTATCTGCTCACTCTTACCGGTTCTCAAGGGGCTCATGGGGGCGGCA 3420
Qy 3117 ACAGGAGCGGACGCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3176
Db 3421 ACCGACCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Qy 3177 TCTGGTTCAGATCGACGCTTTCATCTGTCAGGACCAAGGCGGCGGACGCTGAGGAGTGT 3236
Db 3481 TGTGGTTCAGATCGACGCTTTCATCTGTCAGGACCAAGGCGGCGGACGCTGAGGAGTGT 3540
Qy 3237 GCATCAATTGCT 3248
Db 3541 GCATCAATTGCT 3552
```

RESULT 5

AAZ99494

ID AAZ99494 standard; DNA; 3773 BP.

XX

AC AAZ99494;

XX

```
DT 03-JUL-2000 (first entry)
XX DNA encoding a maize cellulose synthase.
DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
XX transgenic plant; plant breeding marker; ss.
KW Zea mays.
XX
OS Key Location/Qualifiers
FH 337..3565
FT /*tag= a
FT /product= "cellulose synthase"
FT /trans_except= (pos: 1817..1819, aa: Xaa)
FT /note= "no termination codon given; Xaa is an unspecified
XX amino acid"
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
PI P-FSDB; AY84108.
XX
DR WPI: 2000-224343/19.
DR P-FSDB; AY84108.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 90-94; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting or
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 3773 BP; 863 A; 988 C; 1017 G; 904 T; 0 U; 1 Other;
```

Query Match 40.6%; Score 1398.8; DB 3; Length 3773;

Best Local Similarity 67.2%; Pred. No. 3.8e-222;

Matches 2185; Conservative 1; Mismatches 923; Indels 143; Gaps 9;

Qy 133 GAACGAGCTGGTGTGATCCGGGCGCACGAGGACCCCAAGCGCTGGGGCGCTGAGCGG 192

Db 325 GAGCTCGTTCCTGATGGAGGCGGACGCGGCGAGCGGTGAAGTTCGGGAGGCGCGGTGGCGG 384

Qy 193 GCAGGTGTGGAGATATCGCGGACGAGGTCCGGCTCACCGTGCACCGGACCTCTTCGT 252

Db 385 ACAGGTGTGCCAGATCTCGCGCGACGCGGTGGGCGACCCACCGGAGGGGACGCTTCGC 444


```
Db 2604 GCTATGAGCAAGACTGAATGGGAACCTGAGATCGGGTGGATCTACGGTTCTGTGACG 2663
Qy 2337 AGGACATCTCTGACGGGGTTCAAGATGCACTCCGCGGGTGGCGTCCGTGTACTGCAATGC 2396
Db 2664 AAGACATCTCTCACCGGATTCAGATGCAAGCGGAGGCTGCGGTGATCTACTGCAATGC 2723
Qy 2397 CGAAGCGGGCGGGTTCAAGGGGTGCGCGCGCATCAATCTATCGGACCGTCTCAACCAAG 2456
Db 2724 CCAAGCGGCCAGCTTTCAGGGGTCTGCCCCATCAATCTTTTCGAGCCGCTCGAACCAAG 2783
Qy 2457 TGCTCCGGTGGCGCTGGGGTCCGTGAGATCTCTTTCAGCGGCGACAGCCCGCTGTCTGT 2516
Db 2784 TGCTCCGGTGGGTCTTGGGTCCGTGGAGATCTCTTCAGCGGAGCTGCCCGCTGTGGT 2843
Qy 2517 ACGGCTACAAGAACCGCAACCTCAAGTGGCTGGAGCGCTTCGCCCTATCAACACCAACCA 2576
Db 2844 ACGGCTAC--GGAGGGCGGCTCAAGTTCTTGGAGAGATTCGCGTACATCAACACCAACCA 2900
Qy 2577 TCTACCGCTTCACTCGCTCCGCTGCTGGCTTCTGCGCTACTGCAACCTCCCGCGCGTCTGCCCTC 2636
Db 2901 TCTACCGCTTCACTCGCTCCGCTTCTCATCTACTGCAATCTTCCCGCGCATCTGTCTGC 2960
Qy 2637 TCACCGGCAAGTTCATCATGCGCTGATTAGACGTTTCGCGAGCTCTTCTCATCGCC 2696
Db 2961 TCACCGGAAGTTCATCATTCAGAGATCAGAACTTCGCCAGCATCTGGTTCACTCTCC 3020
Qy 2697 TCTTCACTGCTCATCTTCGCGAGCGGCATCTTGGAGATCGCGTGGAGGGGTGAGCATCG 2756
Db 3021 TCTTCACTGCTCATCTTCGCGAGCGGCATCTTGGAGATGAGTGGAGGGGTGAGCATCG 3080
Qy 2757 AGGAGTGTGAGAAAGAGCAGTTCTGGGTCTATCGCGGGGTGTCCGCGCATCTCTTCG 2816
Db 3081 ACAGTGTGTGAGAAAGAGCAGTTCTGGGTGATCGGGGTCATCTCCGCGCACCTCTTCG 3140
Qy 2817 CCGTGTGCGAGGGCTGCTCAAGTCTCTCGCGGATCGACACCACTTCAACCGTCACTT 2876
Db 3141 CCGTGTTCAGGGCTGCTCAAGTGTCTGGCGGATCGACACCACTTCAACCGTCACTT 3200
Qy 2877 CCAAGGCCACCGCGGACGAGACGAGTTTCGCGGAGCTCTACGCTTCAAGTGGACCA 2936
Db 3201 CCAAGGCCCTCG--GACGAGGACGGGACTTCGCGGAGCTGTACATGTTCAAGTGGACCA 3257
Qy 2937 CGTCTCTCATCCCGCCACCAAGCTGTCTCATCTATTAAGTCAATCGGCGTGTGGCGGCA 2996
Db 3258 CGTCTCTGATCCCGCCACCAACCATCTCTGATCATCAACCTGTGGGTGTGTCGCGGCA 3317
Qy 2997 TCTCGGAGCCATCAACAGGCTACAGTCTCTCGGGGCGCTCTTTCGCGAGCTCTTCT 3056
Db 3318 TCTCTTACGCCATCAACAGCGGATACAGTCTGTGGGGCGGCTCTTTCGCGAGCTCTTCT 3377
Qy 3057 TCGCTTCTTGGGTCTATCTCCACCTCTACCGCTTCTCAAGGGGCTCATGGGGCGGCGAGA 3116
Db 3378 TCGCTTCTTGGGTCTATCTCCACCTGTACCGCTTCTCAAGGGGCTCATGGGGCGGCGAGA 3437
Qy 3117 ACAGGAGCCCAACCGTGTGTCTATCTGGTCCATCTGCTGGCTTCCATCTTCTCCCTGC 3176
Db 3438 ACCGACACCCGACCATCTGCTGCTGTGGGCTCTCTGCTGGGCTCTCTTCTCTCTTCT 3497
Qy 3177 TCTGGGTGAGATCGACCTTTCATCTGTCAGGACCAAGGGCGCGGACGTCAGGCAAGTGT 3236
Db 3498 TGTGGGTTCGATCGACCCCTTCCACCCCGCTCACTGGGCGGATACCCAGACGCTGTG 3557
Qy 3237 GCATCAATTTGCT 3248
Db 3558 GCATCAACTGCT 3569
```

RESULT 6
AAA67114
ID AAA67114 standard; DNA; 3851 BP.
XX
AC AAA67114;

```
XX 31-OCT-2000 (first entry)  
XX Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.  
XX Eucalyptus grandis; pinus radiata; Monterey pine; modification;  
XX plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;  
XX transgenic plant; ds.  
XX Pinus radiata.  
XX OS  
XX WO200022092-A2.  
XX 20-APR-2000.  
XX 08-OCT-1999; 99WO-NZ000169.  
XX 13-OCT-1998; 98US-00170862.  
XX 11-AUG-1999; 99US-0148426P.  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX Bloksberg LN;  
XX WPI; 2000-339328/29.  
XX P-PSDB; AAB16307.  
XX New genes encoding proteins involved in a plant polysaccharide  
XX biosynthetic pathway, useful for modulating or altering the  
XX polysaccharide content, composition or structure of the plant.  
XX Claim 1; Page 71-72; 301pp; English.  
XX The present invention describes isolated polynucleotides (PN) comprising  
XX a sequence selected from one of 835 nucleotide sequences given in  
XX AAB67073 to AAB67907, their (reverse) complements, sequences producing an  
XX Expectation (E) value of 0.01 or less compared to the 835 sequences,  
XX sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the  
XX 835 sequences or sequences that are degenerately equivalent or allelic to  
XX the 835 sequences. The polynucleotides are used to modify the activity of  
XX a polypeptide involved in a polysaccharide biosynthetic pathway in the  
XX plant. They are especially used to modulate or alter the polysaccharide  
XX content, composition or structure of the plant. AAB16268 to AAB16340 are  
XX proteins encoded by some of the polynucleotide sequence given in the  
XX present invention  
XX Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 U; 0 Other;
```

```
Query Match 39.7%; Score 1367; DB 3; Length 3851;  
Best Local Similarity 56.5%; Pred. No. 7e-217;  
Matches 2197; Conservative 0; Mismatches 940; Indels 166; Gaps 10;  
Qy 90 CGATGGAGGCCAGCGCGGCTGTGGCGGCTCGGCACACCGGACGAGTGTGTGCTGA 149  
Db 151 CAATGGAAGCCAGCGCGGCTTGTGTCGGTCTCTATAACAGAAACGAGTGTGTGCTCA 210  
Qy 150 TCCGGGCGCCACGAGGACCCCAAGCGCTGCGCGCTGAGCGGCGAGGTGTCCGAGATAT 209  
Db 211 TCATGGACATGAGAGCCGAGCCCTTGAACAGTGTGATGSCCAGCTCTGCCAGATT 270  
Qy 210 GCGGCGACGAGGTGTGGGCTCACGGTGGACCGGACCTCTTTCTGTCGCTCAACGAGTGG 269  
Db 271 GTGCGGAGGACGTCTGGGCTTAAACAGACGCGGAGCTGTTCGTTCCTGTAATGAGTGG 330  
Qy 270 GCTTCCCGGTGTGGCGGCTGTGTACAGTACAGCGCGCGGAGGCGACGAGACTGCC 329  
Db 331 GGTTCCTGTCTGTGGCGGCTGTGTATGAGTACGAGAGACGAGAGGAAATCACTGCTGCC 390  
Qy 330 CCCAGTGCAGACGCGCTACAAGCGCTCAAGGGGAGCCGAGGTTGCCGGGACGATG 389  
Db 391 CGCAGTGCATATCTGTTTACAGCGTCAAAAAGGAGTCCACCGGGTGAAGGTGACGATG 450
```

QY 390 ACGAGGAGACATCGACGACCTGGAGCACGAGTTCAACATCGACGACGAGAAATCAGAGA 449
Db 451 ATGAAGAAGACGTTGATGATGAATGAACATGAATTTTATGTGGAGACTTCAGCAAGAAGAAC 510
QY 450 GGACAGCTGGAGGGCAACATGCAGAAACAGCAGATCACCGAGGCGATGCTGCACGGCAGGA 509
Db 511 GGAG-----CAGATCACCGAGGGGATGCTCCAGGACGCA 546
QY 510 TGAGCTACGGAGGGGCGCGACGCGACCGCAACAAACACCCCGCAGATCCCGCCCA 569
Db 547 TGAGCTATGCCGAGGTCCCGACGACGAAATTCGCAGATTGCTC--ATAATCCAGAGCT 604
QY 570 TCAATCACCGGCTCCCGCTCCGTGC-----CGTGAGCGGTGAGTTCCGAT 615
Db 605 TCCTCCGAGATTCTCTGTACTTGCAACACGGCCACTCGGTTGTGAGTGGGAGATTCCAAC 664
QY 616 TAC---CAACGGGTATGGCCACGCGCAGGTCCTGCTCTCCCTGCACAAGCGCATCCATCC 672
Db 665 GTCACTACGCGACAMACCAATTGCTTGCCAAACCCTGCAATGCTGAAGCGGTGCATCC 724
QY 673 GTACCCCTGTCTTGAGCCAGGAGTGCCAAAGTGGGACGAGAAAG-----719
Db 725 AAGCTCCGAGCGGGAGTGGAGGATCATGTGATCCAAACAGGGATATTGTTCTTA 784
QY 720 -----AAGTGAGCTGGAAGAGAGGATGGACGACTGGAAGTCCAAAGCAGGGCAT 768
Db 785 TGGCTTTGGGAAAGTGTCTTTGGAGGAGCGAGCGATGTTATAAATCGAAGGAAACAA 844
QY 769 CCTCGGCGCGCGCGATCCGAG-----794
Db 845 ATCAGGCCAGTTGGATATGACGAGAGGAGATATCAATATATAATGAGGGGTTTCGACCAAA 904
QY 795 -----ACATGGACGCGCAGCTGGCACTGGAACGAGGAGCGAGGCGCGCT 840
Db 905 TGAGCCTGAAGATTATATTGATCCCGATATGCCAATGACCGATGAAGCAGGCGCACT 964
QY 841 GTCAGGAAGGTGTCGATCCGCTCGACGAAGGTGAACCCGTACCGGANTGTGATCGTGGT 900
Db 965 GTCCGAAAAGTGCRAATTCCTTCAAGCAAAATAAATCCATACCGAATGTCATTGTAAT 1024
QY 901 GCGTCTGTTGTGCTCGCTCTTCTCCTCGGTACCGTATCTCGACCCGTCGCGACGC 960
Db 1025 TCGACTGATAGTCTGGGTATTTTCTCCGCTATCGTCTCTGAAATCCAGTGAAGATGC 1084
QY 961 CATCGGCTGTGGCTCGTCCATCATCTGCGAGATCTGGTTCGCCATCTCTCTGGATCCT 1020
Db 1085 ATATGGGCTCTGGGCCACTTCTATCGTTTGTGAAATCTGTTGGCTTGTGATGATCT 1144
QY 1021 CGACCACTTCCCAAGTGGTTCCCATCGACCGGAGACGTACTCGACCGCTCTCCCT 1080
Db 1145 TGATCACTTCCCAAGTGGTTGCTATCGTCTGTAACGCTATCTGATCGACTGTCAAT 1204
QY 1081 CAGGTACGAGAGGAGGGAGCGCTCGCTGCTGTCGGCGGTGACCTGTTCTGTGAGCAC 1140
Db 1205 AAGTACGAACGAGAAGCGAACATCAATGCTTGCACTGTGACCTCTTGTGAGTAC 1264
QY 1141 GGTGGACCCGCTCAAGGAGCGCGCTGGTGTGACCGCAACACCGTGTCTTCCATCTCTCGC 1200
Db 1265 TGTAGATCCACTGAAGGAGCTCCTTTGGTTACTGCCAATACAGTATTATCAATCCTTTC 1324
QY 1201 CGTAGACTACCCGTGACAAAGTGTCTCTGCTAGCTCTCGACGACGCGCGGTGCATGCT 1260
Db 1325 AGTAGACTACCTGTAGACAAATGTCTCTGTTATGTCTGATGACGAGCGTGCATGCT 1384
QY 1261 GAGCTTCGAGTCTGTGCGAGACGCGCGAGTTTCGCGCAAGTGGTGCCCTTCTGCA 1320
Db 1385 TACTTTTGAATCTCTCTGAGACCTCAGAAATTTGCCAGAAATGSGTACCATTCTGCA 1444
QY 1321 GAAGTTCGGCATCGAGCCCGCGCCCGAGTTCTACTTCTCGCTCAAGTTCGACTACCT 1380
Db 1445 GAAATTGCATTTGAGCTCGCTCGCTCCGAAAATCTATTCTCTCAGAAAATTGACTATCT 1504
QY 1381 CAAGGACAGGTGACGCCCACTTCTGTCGAGGAGCGCGCCCATGAGAGAGATGTA 1440

Db 1505 GAAGGACAAATTTCAACCCACCTTTGTCAAAGAGCGCGTGCATGAAGAGAAATATGA 1564
QY 1441 CGAGTTCAAGGTCCGATCAACGCGCTGTTGGCCAAAGGCATGAAGGTGTCGCGCAGAGG 1500
Db 1565 AGAATTCAGGTGCGCATCAATCGGTTGGTTGCAAGGCTCTAAAGTGCCTCCAGAGAG 1624
QY 1501 GTGGATCATGAAGGACGCGACGCGCTGGCCCGGAAACAACAACCGCGACACACCCCGCAT 1560
Db 1625 ATGGACAAATGCAAGACGGTACGCTTGGCTGCTGTAATAATACCCGTGACCATCTCGTAT 1684
QY 1561 GATCAGGTGTTCTCTGGGCCACAGCGCGGCCACGACGAGGGCAACAGGTGTCGCCCG 1620
Db 1685 GATCCAAGTGTCTTGGGTCACAGTGGCGGCTCGATACAGAAGGCAATGAGCTTCTCTG 1744
QY 1621 CTTCTGTGATGCTCTCCGCTGAGAGCGCGCGGATTTCCAGCACCAAGAAGSCCGCGC 1680
Db 1745 GCTAGTATATGTTTCTGCTGAGAGAGACTGTTTCCAGCATCACAGAGSCCGGTGC 1804
QY 1681 CATGAACGCTCTGATTCGCGTCTCTCGCCGCTGTGACCAACGCGCCCATTCATGCTCAACT 1740
Db 1805 CATGAATGCTTGTGGTTTCGGGTTCTGCTGTGCTCACCAATGCTCCATTTATGCTGAATCT 1864
QY 1741 GACTGTGATCACTACATCAACAAACGAAAGGCCATCCCGGAGGCCATGTGCTTCTCAT 1800
Db 1865 GGATTTGATCACTACATTAACAATAGCAAGGCNAATCAGGGAAGGCATGTGCTTTATGAT 1924
QY 1801 GGAACCTCAGGTGCGCGCGGAAGTCTGTACTGTTCACTGTTCCCGCAGAGGTTCGACGGCAT 1860
Db 1925 GATCTCTCAGTTGGAGAAAGTCTGTTATGTCCAAATTCCTCAGAGATTCGATGAT 1984
QY 1861 CGACGTGACGACCGCATACGCTAAACAGGAACHAACGCTCTTCTTCGACATCAACATGAAGG 1920
Db 1985 TGATCGCAATGACCGTTACGCCAATCGAAACACCGTATTTCTTTGATATCAACATGAAGG 2044
QY 1921 GCTGAGCGCATCAAGGCCCGGTGTACGTCCGGAACAGGGTGGTGTTCGCGCGCACGC 1980
Db 2045 TCTGGAATGGAATTCAGGGCTGTATATGTGGGAACGTGGATGATGTTTCAGAGAACAAGC 2104
QY 1981 GCTCTACGGCTACAAACCTCCCAAGGACCCCAAGAGGCCAAGATGCTGACCTTGCACATG 2040
Db 2105 TCTATATGGGTATGGGCTCCCAAGGCCCAAAAGCTCCCAAGATGCTGATGTTG 2164
QY 2041 CTGCGCGTCTTTCGCGCGCAAGAG-----2065
Db 2165 TCTCCCTTGTGGGCTCTCGTAAGAAAGTCTCCGAAGAAATAATAGTAGCAAGAAAAGTGC 2224
QY 2066 -CGGAACACGCAAGGACGGGCTGCGGAGGCGCACCGCTGATATGGGAGTA-----GA 2118
Db 2225 AGGAATCCAGCTCCCGCTACAATCTGACCGGATTCGAGGAAGGATGAAGGTTATGA 2284
QY 2119 TAGCGACAAAGGAGATGCTCATGTCCCATGAATTCGAGAAAGCGGTTCCGGCAGTCCCG 2178
Db 2285 TGACGAAGAGCATGTTGATGAGCCACTAGACTTCGAGAGAGAGTTTGGCCAGTCTTC 2344
QY 2179 GCGTTTGTGTCAGCTCGACGCTGATGGAGAAAGCGCGCTCTCTCTTCTGTCGAGCCCCG 2238
Db 2345 AGCTTTTGTTCATCCACTCTGATGGAGAAATGTTGGTGTTCGCAAAACAGCAATCCAGC 2404
QY 2239 CGGCTCTCAAGGAGGCCATCCATGTCTCATGCTGCGGCTACGAGGACAGACCCGACTG 2298
Db 2405 TGAATTTGTTGAAGGAGGCTATTCAATGTCTCATGCTGTGGATATGAAGACAAAACGGAATG 2464
QY 2299 GCGGCTGAGAGTCTGGGTTGGATCTACGCGGTGATCACGAGGACATCCTCAGCGGTTCAA 2358
Db 2465 GGGAAAGAGCTTGGATGATCTATGGAATCAGTCAAGAGGACATTTCTGATGGATCAA 2524
QY 2359 GATGACTCCCGCGGTGCGCTCCGTGTACTGCAATGCGAAGCGGCGCGGCTTCAAGGG 2418
Db 2525 GATGCACTCGAGGCTGGCGGTCCATTTACTGTATGCCCAACAGCAGACGATTCAGAGG 2584
QY 2419 GTGCGCGCGATCAATCTATCGAACCGTCTCAACAGGTGCTCCGCTGCGCGCTGGGTC 2478

```
Db 2585 GTCTGCTCCAATCAATCTATCAGACCGTTTGAAACAGAGTGTGGTGGGCTTTGGGATC 2644
Qy 2479 CGTCGAGATCTTTCACGCGGCAAGCCCCCTCTGTACCGCTACAGAAACGCAACCT 2538
Db 2645 AGTAGAAATTTTCATGACGACATATGCCAATCTGGTATGGCTATGGG---GGAGGTCT 2701
Qy 2539 CAAGTGGCTGAGCGCTTCGCTACATCAACACACCACTACCCCTACCTCGCTCC 2598
Db 2702 GAAATGGCTTGAAGAATTTGGCTATATCAACACCAATGTCTATCAATTCACCTCTCTCC 2761
Qy 2599 GCTGCTGCTCTACTGCACTCCCGCGCTCTGCTCTCACCAGGCAAGTTTCATCATGCC 2658
Db 2762 ACTCATTTGGCTATTCGACACTTCAGCGCTCAGTTTGTCTCACTGGCAATTTGTATCCC 2821
Qy 2659 GTCGATAGACAGTTCGCGACGCTTCTTCTCATCCGCTCTTCATGTCATCTTCGCGAC 2718
Db 2822 TCAGATCAGTACTTTTGGAGCTATTTTAAATAGCTCTTTTATCATCAATTTTGGCAC 2881
Qy 2719 GGGCATCTGAGATCGGCTGGAGCGGGGTGAGCATCGAGAGTGGTGGAGAACGAGCA 2778
Db 2882 TGGTAATCTGAAATGAGGTGGAGTGGAGTGAATGGAATGGTGGCGAAATGAACA 2941
Qy 2779 GTTCTGGGTCTATCGCGCGGTTCGCGCATCTCTTGCCGCTCTGTCAGGSCCTGCTCAA 2838
Db 2942 GTTCTGGGTATTCGAGGGGTTTCTGCACATTTTTCAGTATTAAGTCTGCTCAA 3001
Qy 2839 GGTCTCGCGGGATCGACACCAACTTCACGCTCACCTTCAAGGCCACCGCGCAGGGA 2898
Db 3002 GGTACTGCGAGGCATTTGATACAAATTTTACACAGTCACTGCCAAGGCATC-----AGATGA 3055
Qy 2899 CGAGAGTTGCGGAGCTCTAGCGCTTCAAGTGAACGCTCTCTCATCCGCGCCACAC 2958
Db 3056 CGGTGAGTTTGGGAACGTATGATCAATCAAAATGACACCACTCTCTCATTTCTCTACAAC 3115
Qy 2959 GCTGCTCATATTAACGCTCATCGCGCTCGTGGCGGATCTCCGACGCAATCAACAACGG 3018
Db 3116 CTTGCTGTCTATCAACCTTTGGGGGTGTTTGTGGGTAGCAGATGAATCAACAATGG 3175
Qy 3019 GTACAGTCTGGGGGCCCCCTTTCCGCAAGCTCTTCTTCCGCTTCTGGGTCTATCGTCCA 3078
Db 3176 ATTTTCAGTCAATGGGTCTCTCTTTGGGTAAAGCTTTCTTTGCAATCTGGGTCAATGTGCA 3235
Qy 3079 CCTTACCCGTTCTTCAAGGGCTCATGGGGCGGCAGAAAGGACGCCACCGTTGTGT 3138
Db 3236 CTTGTATCTTCTTCTCAAGGGTCTCATGGGCGAGGACGAACCGAACCCACCATCTGTGT 3295
Qy 3139 CATCTGTTCCATTTCTGTGGCTCCATCTTCTCTCTGCTCTGCTGAGGATCGACCTTT 3198
Db 3296 TATTTGGTCAATTTCTGTGGCATCTGTTTCTCTTCTTCTGGGTAAGAATTTGATCTTT 3355
Qy 3199 CATCGTCAGGACCAAGGGCCCGGACGTCAGGAGTGTGGCATCAATTTGCTGAGCTGTTA 3258
Db 3356 CTTGAGTAAGGTTAAAGGCCCAGATACTAAACAATGTGGCATCAACTGCTGATTTCTTTG 3415
Qy 3259 TTA 3261
Db 3416 ATA 3418
```

RESULT 7

AAZ58265

ID AAZ58265 standard; cDNA; 3786 BP.

XX

AC AAZ58265;

XX

DT 08-MAY-2000 (first entry)

XX

DE Corn cellulose synthase cDNA clone p0097.cqradi7rc.

XX

KW Corn; maize; cellulose synthase; transgenic plant; ss.

XX

OS Zea mays.

XX

FH Key Location/Qualifiers

CDS 2..3499

FT /*tag= a

FT /partial

XX

PN WO200004166-A2.

XX

PD 27-JAN-2000.

XX

PF 13-JUL-1999; 99WO-US015871.

XX

PR 14-JUL-1998; 98US-0092844P.

XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX

PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;

PI Rafalski JA, Thorpe CJ;

XX

XX

DR WPI; 2000-182431/16.

DR P-PSDB; AAY58834.

XX

PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as

PT probes for isolating cDNAs and genes encoding homologous proteins, for

PT producing transgenic plants.

XX

PS Claim 9; Page 44-45; 93pp; English.

XX

CC The present sequence is that of cDNA clone p0097.cqradi7rc encoding a

CC portion (see AAY58833) of corn cellulose synthase (CS). The cDNA clone

CC was isolated from a European corn borer 4 times-infected corn stage V9

CC whorl section cDNA library on the basis of homology to Arabidopsis and

CC cotton CS sequences. The invention relates to isolated nucleic acid

CC fragments encoding plant CS and to CS polypeptides. It also relates to

CC the construction of a chimeric gene encoding all or a portion of the CS,

CC in sense or antisense orientation, where expression of the gene results

CC in altered levels of the CS in transformed host cells. The host cells can

CC be used to screen compounds for their ability to inhibit CS activity. CS

CC nucleic acids are also useful for producing transgenic plants having

CC altered levels of CS, and hence altered levels of fibre. CS may also

CC serve as a target for the development of novel herbicides

XX

SQ Sequence 3786 BP; 846 A; 967 C; 1111 G; 862 T; 0 U; 0 Other;

Query Match

Best Local Similarity 39.6%; Score 1364.8; DB 3; Length 3786;

Matches 2202; Conservative 65.7%; Pred. No. 1.6e-216;

Matches 2202; Indels 143; Gaps 9;

Qy 30 CTGTGAGCACCTGAGGGTCCGAGGCGGAGAGTACCTAGCACGCGGCTCCGCGC- 88

Db 151 CTGTGCGCGCGCGCGGGTTCGTGCGCGAGAGATCCGCGGGCGGGCGGGGCGCC 210

Qy 89 -GCGATGGAGGCCAGCGCGGGTGGTGGCGGGTTCGCACAAACCGGAACGAGCTGGTCT 147

Db 211 TGAGATGGAGGCTAGCGCGGGTGGTGGCGGGTTCGCATTAACCGGAACGAGCTGGT 270

Qy 148 GATCCGGGGCCA-CGAGGACCCCAAGCGCTGCGGGCGG-----CTGACGGGCA 195

Db 271 GATCCGGCGGACCGCGAGTCCGAGGCGCGGGCGGGCGCGCGCGCGCGGAGGC 330

Qy 196 GGTGTGGAGATATGCGGCGAGGTCGGGGTTCACGGTGGAGCGGCGACCTTCTGCTGC 255

Db 331 GCCGTGCGAGATATGCGGCGAGGTCGGGGTGGGGTTCGACGGGGAGGCTTCTGTTGC 390

Qy 256 CTGCAACGAGTGGGCTTCCCGGTGTCGGCGCTCTGTACAGTACGAGCGCGGGAGGG 315

Db 391 GTGCAACGAGTGGGCTTCCCGGTGTCGGCGCTCTGTACAGTACGAGCGCGGGAGGG 450

Qy 316 CAGCGAAGTGGCCCCAGTGAAGCGCGCTAACAGCGCTCAAGGGGAGCCCGAGGGT 375

Db 451 CTGCAAGCGTGGCGCGAGTGCAGGACCCGCTCAAGCGCTCAAGGGTGGCGCGGGT 510

Qy 376 TGCGGGGAGCATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 435

Db 511 GCGCGGCGACGAGGAGGAGCGCGCTCGACGACCTGGAGGGCGAGTTCCGCCCTCGAGGA 570
Qy 436 CGAGAAATCAGCAGAGGCGAGCTGAGGGCAACATGTCAGAAACAGCAGATCACCAGAGCGAT 495
Db 571 CGCGCGCGCCACGAGGACGACCGGAGTACGTGCGCGAGTCCATGCTCAGGCGCGAGAT 630
Qy 496 GCTGCGAGGAGGATGAGTACGCGGAGGGGCCCGGACGACGCGGACGCGCAACACCC 555
Db 631 GAGCTACGCGCGCGGCGGCGACGCGACCCCGGCTTACGCCCCGTCCCCAACGTGCGGCT 690
Qy 556 GCAGATCCCGCCCATCATCACCGGCTCCCGCTCGGTGCGGTGAGGTGAGTTCGGAT 615
Db 691 CCTCACCAACGCGCAGATGTTGATGACATCCCGCGGAGCAGACGCGCTCGTGCCTC 750
Qy 616 TACCAACGGGTATGGGCAACGCGGAGGTCTCGTCTTCCCTGCAACGCGCATCCATCCGTA 675
Db 751 CTACATGAGCGGCGGCGGCGGCGGCGGCAAGAGATCCACCGCTCCCTTTCGAGATCC 810
Qy 676 CCTGTGTCGACGCGGAGGTGCCAAGTGGGACGAGAAGAA----- 718
Db 811 CAACCTTCCAGTGCACCCAGATCCATGGACCGGTCCAAGGATCTGCGCGCTTACGGATA 870
Qy 719 -----GAAGTCAGCTGGNAGAGAGGATGGACGACTGGAGTCCAAAGCGGCATCCT--- 771
Db 871 TGGCAGCGTGGCTTGAAGAGAGAAATGGAGGGCTGGAGCAGAGAGAGCGCCTGCA 930
Qy 772 -----CGGCGGCGGCGCGCATCCGGAAGACATGGACCGCGCATGGGCACTGAA 819
Db 931 GCATGTCAGAGCGAGGTGGCGGTGATTTGGGATGGCGAGTGCAGATCTGCCATAT 990
Qy 820 CGACGAGGCGAGGCGCGCTGTCGAGGAAGGTGTGATCGCTCGAGCAAGGTGAACCC 879
Db 991 GGAATGAAGCTAGGCGAGCATTGTCCAGAAAGTCCCTATATCATCAAGCCGAATTAATCC 1050
Qy 880 GTACCGGATGATCGTGGTGGCTCTCGTGTGCTCGCTTCTTCTCGGTACCGTAT 939
Db 1051 CTACAGGATGATATCGTTATCCGGTTGGGTGTTTGGGTGTTCTTTCACCTACCGAGT 1110
Qy 940 CTTGCAACCCCGTCCCGAGCCCATCGGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTG 999
Db 1111 GATGATCCCGGCAAGATGCAATTGCAATTGTGGCTCATATCTGTAATCTGTGAAATCTG 1170
Qy 1000 GTTCGCCATCTCTGAGATCTCGACCAAGTTCCTCCAAAGTGTTCCTCCATCGACCGCGAGAC 1059
Db 1171 GTTTGGCATGCTCTGGATTCTGATCAGTTTCCAAAGTGGCTTCCAATCGAGAGAGAGAC 1230
Qy 1060 GTACCTCGACCGCTCTCCCTCAGGTACGAGAGGAGGGAGCGGCTCGCTGCTGCTGCGC 1119
Db 1231 TTACCTGGACCGTTTGTCACTAAGTTTGAAGGAGGTTCAACCTCTCTCAGCTTGTCTC 1290
Qy 1120 GGTGGACCTGTTGTCGACACGCTGGACCCGCTCAAGGAGCCGCGTGGTGACCCGCAA 1179
Db 1291 AATCGACTTCTTGTGATGAGGTTGATCCACAAAGGAACTCCCTTGTTCACAGGAA 1350
Qy 1180 CACGCTGCTCTCATCTCGCGGTAGACTACCCGTTGGGAACAAGTCTCTGCTACGTCCTC 1239
Db 1351 CACTGCTCTTCCATCTCTGTTGATTTATCCGTTTGAAGGTTCTCTGCTATGTTTC 1410
Qy 1240 CGACGAGCGCGCTCGATGCTGACGTTTCGATCGCTGTCGAGACGCGCCAGTTCCGGC 1299
Db 1411 TGATGATGGTGTGCAATGCTTACGTTTGAAGCATTTGCTGAAACATCTGAAATTTGCAA 1470
Qy 1300 CAAGTGGGTGCTTCTGCAAGAAAGTTTCGGCATCGAGCCCCCGCGGAGTTCTACTT 1359
Db 1471 GAAATGGTTCCTTTCAGCAAAAAGTTTAATATCGAGCTCGTGTCTCTGAGTGTACTT 1530
Qy 1360 CTCGCTCAAGGTGCACTACTCAAGGACAAGGTGCGAGCCCACTTCTGTGCAAGCGCGC 1419
Db 1531 CCAACAGAAATAGACTACCTGAAAGACAAGGTTGCTGCTTCATTTGTTAGGAGAGGAG 1590
Qy 1420 CGCCATGAAGAGAGATGATGAGGTTTCAAGTCCGATCCGATCAACGCGTGTGTCGCAAGC 1479
Db 1591 GCGCATGAAGAGAGAAATACGAGGAAATTCAGGTTAAGGATCAATGCTTGTGTTGCAAAAGC 1650

Qy 1480 CATGAAGGTGCGGCGAGAGGGGTGGATCATGAAGGACGCGACCGCTGGCCCGGGGAACAA 1539
Db 1651 CCAAAGGTTCTCTGAGGAAGGATGCAATGCAAGATGGAAGCCCTGCGCTTGGAAACAA 1710
Qy 1540 CACCGCGACACACCCCGGATGATCCAGGTGTTCTTGGGCCACAGCGGCGGCCACACAC 1599
Db 1711 CGTACGCGATCATCTCTGGAATGATTCAGGTATTCCTTGGCCAAAGTGGCGGTCTGTATGT 1770
Qy 1600 CGAGGSCAACGAGCTGCCCGCTCGTGTACGTCTCCCGTGAGAGCGCCCGGATTCCTCA 1659
Db 1771 GGAAGAAATGAGTTGCTCGCTGCTGTTTATGTTCTCGAGAGAAAGAGGCCAGGTTATAA 1830
Qy 1660 GCACCAACAAGAGCGCGCCCATGAACGCTCTGATTCGCTCTCCCGCTGTGTACCAA 1719
Db 1831 CCATCACAAAGAGGCTGGTGCCATGAATGCACTGGTCCGTGCTCTGCTCTTATCAAA 1890
Qy 1720 CGCGCCATTCATGCTCACTTGGACTGTGATCACTACATCAACACAGCAAGCCATCCG 1779
Db 1891 TGTGTCATACCTATTTGAACCTTGGACTGTGATCACTACATCAACATAGCAAGGCCATAA 1950
Qy 1780 GGAGGCGATGTCTTCTCATGGAACCTCAGGTCCGCGGAAAGGTCTGTACGTTCAAGTT 1839
Db 1951 AGAGGCTATGTGTTTCATGATGATCTTTGGTGGGAAAGTGTGTATGTACAGTT 2010
Qy 1840 CCGCGAGAGTTTCGACGGCATCGAGTGCACGACGATACGCTAACAGGAACAACCGTCTT 1899
Db 2011 CCTCAGAGGTTTGATGGTATTTGACAAAATGATCGATACGCTAACAGGAACGTTGCTT 2070
Qy 1900 CTTTCGACATCAACAGAGGGCTGACCGGATCCAGGCCCGGTGTAGCTCGGACAGG 1959
Db 2071 TTTTTCATCAACATGAAAGTTTGGACGGTATTCAGGACCCATTTATTTGGGTACTGG 2130
Qy 1960 GTGCGTGTTCGCGCGCAGCGCTCTACGGCTACAACTCCCAAGGAGCCCAAGAGGCC 2019
Db 2131 ATGTGTTTTCAGAGCGCAGCACTGTATGTTTATGATGCTCTTAAACGAGAGGCCACC 2190
Qy 2020 -----CAAGATGGTGAACCTGCGACTGCTGCCGTGCTTCGGCGCGCAAGAGCG 2067
Db 2191 ATCAAGAACTTGCACCTGCGCCCAAGTGGTGCCTCTCTGCTGTCGACAGCAAGAACAA 2250
Qy 2068 GAAACAGCGCAAGGACGGGCTGCGGAG----- 2095
Db 2251 GAATAAAGAGAGACTACAAAACCAAGAGCGAGAGAAAGAAATATTTTTCAGAA 2310
Qy 2096 -----GGCACCGCTGATATGGAGT 2115
Db 2311 AGCAGAAACCCATCTCTGTCATATGCTTTGGGTGAAATTTGATGAAGGTGCTCCAGGTGC 2370
Qy 2116 AGATAGCGACAAGAGAGATGCTCATGTCCACATGAATCTTCGAGAAAGCGTTTCGGGAGTC 2175
Db 2371 TGATATCGAGAAAGCGCGAATCGTAAATCAACAGAAATAGAGAGAAATTTGGGCGATC 2430
Qy 2176 CGCGGCTTCGTTCACGTCGACGCTGATGAGGAGAGCGCGGCTCCCTCTCTGCTGAGCCC 2235
Db 2431 TTTCTGTTTTGTGTCATCAACACTTCTTGAGAACGAGGAGCCCTGAAAGCGCAAGTCC 2490
Qy 2236 CGCGCGCTCTCAAGGAGGCCATCCATGTCTACGTGCGGTACGAGGACGAAGACCGA 2295
Db 2491 AGCTTCTCTTGAAGGAAGCTATACATGTTATCAGCTGCGGCTACGAGACGAAGACCGA 2550
Qy 2296 CTGGGGCTGAGAGCTGGGTGGATCTACGGGTGATCAAGGAGGACATCTGACGGGGTT 2355
Db 2551 CTGGGGAAGAGATGCTGCTGATTTACGGATCGATCACAGGATATCTTGACTGAT 2610
Qy 2356 CAAGATGCACTCGCGGGTGGGCTCCGTGTACTGATGCCGAGGAGGGCGGCTCAA 2415
Db 2611 TAAGATGCACTGCCATGGCTGGCGGTCTATTTACTGCATCCCGAAGCGGCTGCAATCAA 2670
Qy 2416 GGGGTGCGCGCGCATCAATCTTACGAGCGGTCTCAACACAGGTGCTCCGGTGGCGGCTGGG 2475
Db 2671 AGGTTCTGCGCTCTGAACTTTCCAGCGCTCTTCAACAGGTCTTTCGCTGGGCGCTTGG 2730

Qy 2476 GTCCGTCGAGATCTTCTTTCAGCGGCACAGCCCTCTGTACGGCTACAGAACGGCAA 2535
Db |||||
Qy 2731 GTCCGTCGAAATTTCTTTCAGCAAGCACTGCCCATTTGGTACGGATAC---GGCGCGG 2787
Db |||||
Qy 2536 CCTCAAGTGGCTGAGCGCTTGGCTATACCAACACACATCTACCCCTTCACTCGCT 2595
Db |||||
Qy 2788 GCTAAATTCCTGGAAGGTTTCTTATATCACTCCATCGTTTATCCCTGGAGCTCAT 2847
Db |||||
Qy 2596 CCCTCTCTGCTGCTTACCTTCCCTCCCGCGCTTGTCTCTACCGGCAAGTTTCATCAT 2655
Db |||||
Qy 2848 TCCTCTCTGCTTACTTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2907
Db |||||
Qy 2656 GCCTGATATGAGCGTTCGCGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2715
Db |||||
Qy 2908 ACCAGAGCTTACCAATGTCGCGAGTATCTGGTCTTCTGCTTCTTCTTCTTCTTCTTCT 2967
Db |||||
Qy 2716 GACGGGATCTCTGAGATGCGTGGAGCGGGTGGAGCATCGAGAGTGGTGGAGAACGA 2775
Db |||||
Qy 2968 GACCGGATCTTGAATGAGTGGAGTGGCTGGCCATCGACGACTGGTGGAGAACGA 3027
Db |||||
Qy 2776 GCAATCTGCTGCTATCGCGCGCTGCTCGCGCATCTTCTTCTGCGCTGCTGAGGCGCTGCT 2835
Db |||||
Qy 3028 GCAATCTGCTGCTATCGAGCGCTTTCGGGCGATCTGTTGCGGCTGTTCCAGGCGCTGCT 3087
Db |||||
Qy 2836 CAGGTCTCTCGCGGATCGACACCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2895
Db |||||
Qy 3088 GAAGGTGTTTCGCGGCTGCGACGAGCTTTCACGCTGCGAGCGTGGAGCGGAGCGACGA 3147
Db |||||
Qy 2896 GGACGAGCTTTCGCGGCTGCTGAGCTTTCAGGCTTTCAGGCTGCGACGCTGCTGCTGCTGCT 2955
Db |||||
Qy 3148 G-----GAGTCTCGAGCTGCTGACGTTCAAGTGGACACGCTGCTGCTGCTGCTGCTGCT 3201
Db |||||
Qy 2956 CAGCTCTCATCATTAACGTCATCGCGCTGCTGCGCGCATCTTCCGACGCGCATCAACAA 3015
Db |||||
Qy 3202 CAGCTCTCTCTGCTGAATTCATCGCGGCTGCTGCGGATCTTCCGAACGCGATCAACAA 3261
Db |||||
Qy 3016 CAGGTACAGTCTCGCGGCTGCTGCGGAGCTTCTTCCGAGCTTCTTCCGCTTCTGCTGCTGCT 3075
Db |||||
Qy 3262 CAGGTACAGTCTCGCGGCTGCTGCGGAGCTTCTTCCGAGCTTCTTCCGCTTCTGCTGCTGCT 3321
Db |||||
Qy 3076 CCACCTCTACCGCTTCTCAAGGCTGCTGCGGCTGCTGCGGAGCTTCTTCCGAGCTTCTTCCG 3135
Db |||||
Qy 3322 CCACCTGACCGCTTCTCAAGGCTGCTGCGGAGCTTCTTCCGAGCTTCTTCCGAGCTTCTTCCG 3381
Db |||||
Qy 3136 TGTCATCTGCTTCTTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3195
Db |||||
Qy 3382 CATGCTCTGCTGCTTCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3441
Db |||||
Qy 3196 TTTTCATCTGAGGACGAGGCGCGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAG 3247
Db |||||
Qy 3442 GTTCTCTGCGCAAGAGCAACGCGCGCTCTTGGAGGAGTGTGGCTGGAAGTGC 3493
Db |||||

RESULT 8

AZ99509
ID AZ99509 standard; DNA; 3813 BP.

XX
AC AZ99509;

XX 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.

XX Zea mays.

XX Key Location/Qualifiers
FH 215..3493
CDS /*tag= a

FT /product= "cellulose synthase"
FT /note= "no termination codon given"

XX W0200009706-A2.
PN
PD 24-FEB-2000.
XX 16-AUG-1999; 99WO-US018760.
XX 17-AUG-1998; 98US-0096822P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Dhugga KS, Helenjaris TG, Bowen BA, Wang X;
XX WPI; 2000-224343/19.
XX P-PSDB; AAY84113.
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
XX Claim 1; Page 129-134; 119pp; English.
XX The present sequence encodes a maize cellulose synthase polypeptide. The
XX cellulose synthase can be used for the improvement of stalk quality for
XX improved stand or silage. It also provides an increased concentration of
XX cellulose in the pericarp, hardening the kernel and improving its
XX handling ability. The sequences are used to produce transgenic plants and
XX seeds expressing the cellulose synthase. The polynucleotide is used for
XX modulating, preferably increasing, the level of the synthase in a plant
XX cell. The plants are preferably monocots. The polynucleotide is also used
XX as a probe or primer in the detection quantitation or isolation of gene
XX transcripts. The probes are useful in detecting deficiencies in the level
XX of mRNA in screenings for desired transgenic plant, for detecting or
XX mutations in the gene, for monitoring upregulation of expression or
XX changes in enzyme activity in screening assays of compounds, for
XX detection of any number of allelic variants of the gene, or for use as
XX molecular markers in plant breeding programs. The isolated nucleic acids
XX of their encoded polypeptides or for use as immunogens in the preparation
XX and/or screening of antibodies. The proteins can be employed in assays
XX for enzyme agonists or antagonists of enzyme function or for use of
XX immunogens or antigens to obtain antibodies specifically immunoreactive
XX with a protein

XX SQ Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 U; 0 Other;

Query Match 39.6%; Score 1364.8; DB 3; Length 3813;
Best Local Similarity 65.7%; Pred. No. 1.6e-216;
Matches 2202; Conservative 0; Mismatches 1007; Indels 143; Gaps 9;

Qy 30 CTGTCGAGCACTGAGGGGTGCGAGGCGGAGAGCTAGCTTAGCAGCGCGGCTCCGCGC- 88
Db 151 CTGTGCGCGCGCGCGGGTTCGTGCGAGGAGATCCGCGGGCGGGCGGGGCGC 210
Qy 89 -GGCATGGAGGCGCAGCGCGGGTGTGGCGGCTGCGACAAACCGGACGAGCTGTGCT 147
Db 211 TGAGATGGAGGCTAGCGCGGGTGTGGCGGCTGCGATACCGGACGAGCTGTGCT 270
Qy 148 GATCGGGGCGCA-CGAGGACCCCAAGCGCTGCGGGCG-----CTGAGCGGGCA 195
Db 271 GATCGCGCGGACCGCGAGTTCGGGAGCGCGGGCGGGCGGGCGGGCGGGGAGGC 330
Qy 196 GGTGTGCGAGATATGCGGCGACGAGGTTCGGGCTCAGCGTGGAGCGGCGACCTTCTGTCGC 255
Db 331 GCCGTGCCAGATATGCGGCGACGAGGTTCGGGTGGGCTTCGACGGGAGCCCTTCGTGGC 390
Qy 256 CTCGAACGAGTGGGCTTCCCGTGTGCGGCGCTGCTAGAGTACGAGCGCGGGAGGG 315
Db 391 GTGCAACGAGTGGGCTTCCCGTGTGCGGCGCTGCTAGAGTACGAGCGCGGGAGGG 450
Qy 316 CAGCGAGACTGCGCGGCGGCTTCAAGCGCTTCAAGCGCTTCAAGCGCTTCAAGCGCTTCAAGCG 375
Db 451 CTCGCAAGCTGCGCGGCGGCTTCAAGCGCTTCAAGCGCTTCAAGCGCTTCAAGCGCTTCAAGCG 510

Qy	376	TGCGGGGACGATGACGAGGAGCAATCGACGACTGGAGCAGAGTTCAACATCGACGA	435
Db	511	GGCGGGGACGAGGAGGAGACGGCGCTCGACGACCTGGAGGCGAGTTTGGCGCTCGAGGA	570
Qy	436	CGAATATCAGCAGAGCAGCTGGAGGGCAACATGCAGACGACGACGATCAGGAGCGAT	495
Db	571	CGGGCGGGCCACGAGGACGACCCGAGTACGTTCGCGGAGTCCATGCTCAGGGGCGAGAT	630
Qy	496	GCTGCACGGCAGGATGAGCTTACGGGAGGGGCGCCGACACGACGCGACGGCAACACACCCC	555
Db	631	GAGCTACGGCGGGGGCGACGCGCACCCGGCTTCAGCCCGTCCCCACAGCTGCCGCT	690
Qy	556	GCAGATCCCGCCCATCATCAGCGGCTCCCGCTCGGTGCGGTGAGCGGTGAGTTTCGAT	615
Db	691	CCTCACCAACGGCCAGATGGTTGATGACATCCCGCGGAGCAGCAGCGGCTCGTGGCGGTC	750
Qy	616	TACCAAGGTATGGCCACGGCGAGTCTCGTCTTCTTCCCTGACACAGCGCATCCATCCGTA	675
Db	751	CTACATAGCGGGGGGGGGGGGGGGAAGAGATCCACCCGCTCCCTTTTCGAGATCC	810
Qy	676	CCCTGTGTCTGACCCAGGAGTGCCTAAGTGGGACGAGAGAAA-----	718
Db	811	CAACCTTCCAGTGCACCGAGATCATGACCCGCTCCAGGATCTGGCGCGCTACGGATA	870
Qy	719	----GAAGTGAGCTGAAAGGAGGATGGACGACTGGAAGTCCAAAGCAGGGCATCCT---	771
Db	871	TGGCAGCGTGGCTCGAAGGAGAGATGAGGGCTTGAAGCAGAGCAGGAGCGCTGCA	930
Qy	772	-----CGGCGGGCGGCGCATCCCAAGACATGACGCGCGAGCTGGCATGAA	819
Db	931	GCATGTACAGAGCGAGGGGTGGCGGTGATTTGGGATGGCGACGATCTGCCACTAAT	990
Qy	820	CGACGAGGCGAGCAGCGCTGTGCAGGAAGGTGTCGATCGCGTCGACGAAGGTGAACCC	879
Db	991	GGATGAGCTTAGCGAGCCATTGTTCAGAAAAGTCCCTATATCATCAAGCCGAAATTAATCC	1050
Qy	880	GTACCGGATGGTGATCGTGCGTCTCGTTGTGTGCTCGCCTTCTTCCTCGGTAACCGTAT	939
Db	1051	CTACAGGATGATTATCGTTATCCGTTGGTGTGTTTGGGTTCCTTCTCCACTACCGAGT	1110
Qy	940	CCTGCACCCCGTCCCGGACGCCATCGGGCTGTGGCTGTCTGCATCATCTCGAGATCTG	999
Db	1111	GATGCAATCCCGGAAAGATGCAATTTGCAATTTGGCTCATATCTGTAATCTGTGAATCTG	1170
Qy	1000	GTTCCGCATCTCTGGATCTCGACAGTTCGCCAAGTGGTTCCCATCGACCGCGAGAC	1059
Db	1171	GTTTGCATGCTCTGATTTCTGATCAGTTTCCCAAAGTGCGTTCCAATCGAGAGAGAC	1230
Qy	1060	GTACCTCGACCGGCTCTCCCTCAGGTAACGAGGGAAGGGAGCGCTGCTGTCTCGGC	1119
Db	1231	TTACCTGGACCGTTTGTCTAAGGTTTGACAAAGGAGGTCAACCTCTCAGCTTGCTCC	1290
Qy	1120	GGTGGACCTGTTCTGAGACAGGTGACCCGCTCAAGGAGCGCGCTGGTGACGCCAA	1179
Db	1291	AATCGACTCTTTGTTCAGTACGGTTGATCCCAAGAGGAACCTCCCTGGTTCACAGGAA	1350
Qy	1180	CACGCTCTCCATCCTCGCGTAGACTACCCCGTGGACAGGTCTCCTCTAGCTCTC	1239
Db	1351	CATGTCTTTCCATCCTTCTGTGATTTATCCGTTGAGAAGGTCTCCTCTATGTTTC	1410
Qy	1240	CGACGACGGCGCTCGATGCTGACGTTTCAGTTCGCTGTGAGAGCGGCGAGTTTCGGCG	1299
Db	1411	TGATGATGTGCTGCAATGCTTACGTTTGAAGCAATGTCTGAAACATCTGAATTTGAAA	1470
Qy	1300	CAAGTGGTGCCCTTCTGCAAGAGTTTGGCATTCGAGCCCGCGCCCGGAGTTCTACTT	1359
Db	1471	GAATGGGTCTCTTTCAGCAAAAAGTTTAATATCGAGCTCGTGCTCGTGAAGTGTACTT	1530
Qy	1360	CTCGCTCAGGTCGACTACTCAAGCAGAGGTGACGCCACCTTCCTGCGAGGAGCGCG	1419
Db	1531	CCAAAGAGATAGACTACTGAAAGACAAGGTGTGCTCTCATTTGTTTGAAGGAGAGAG	1590

QY 316 CAGCGAGAACTGCCCCAGTGCAGAGCGGCTACAGCGGCTCAAGGGAGCCGAGGTT 375
Db 451 CTCGCAAGCTGCCCCAGTGCAGAGAGCCCGCTCAAGCGCTCAAGGGCTGCCCCGGGT 510
QY 376 TGC CGGGGACGATGACGAGGAGGACATCGAGCGACCTGGAGCAGAGTTCAACATCGACGA 435
Db 511 GGC CGGGGACGAGGAGGAGAGCGGCTGAGAGCTGGAGGGGAGTTGCGCTTGAGGA 570
QY 436 CGAGAAATCAGCAGAGGAGCTGGAGGGCAACATGTCAGAAACAGCGAGATCAACGAGGCGAT 495
Db 571 CGCGCGCGCCACGAGGACGACCGCAGTAGCTGCGCGAGTCCATGCTCAGGCGCGAGAT 630
QY 496 GCTGCGCGCAGAGTAGCTACGGGAGGGGCCCCGAGCAGCGGCAACGCGCAACACACCCC 555
Db 631 GAGCTACGGGCGGGCGGCGAGCGCACCCCGGCTTCAGCGCGCTCCCAACGTGCGGCT 690
QY 556 GCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGCTGCGGTCAGCGGTGAGTTCCGAT 615
Db 691 CCTCACCAACGGCCAGATGTTGATGACATCCCGCGGAGCAGACCGGCTCGTGCCTC 750
QY 616 TACCAACGGGTATGGCCACGCGGAGGTTCTCGTCTTCCCTGCAACAGCGCATCCATCCGTA 675
Db 751 CTACATGAGCGGCGGCGGCGGGGCAAGAGGATCCACCGCTCCCTTTGCGAGATCC 810
QY 676 CCCTGTGTCTGAGCCAGGAGTGCCCAAGTGGGAGGAGAAA----- 718
Db 811 CAACCTTCCAGTGCAACCGAGATCCATGGACCGCTCCAAGGATCTGCGCGCTACGGATA 870
QY 719 ----GAAGTCAGCTGGAGGAGAGGATGGACGACTGGAGTCCAAGCAGGGCATCCT--- 771
Db 871 TGGCAGGCTGGCTGGAGGAGAGATGGAGGGCTGGAGCGCTGGAGCAGGAGCGCTGCA 930
QY 772 -----CGGCGGCGGCGCGATCCCGAGAGCATGGAGCCGCGAGCTGGCACTGAA 819
Db 931 GCATGTGAGGAGCGAGGCTGGCGTGATTTGGGATGGCGAGATGCGACACTGCTCACTAAT 990
QY 820 CGAGGAGGAGGAGCGGCTGTGAGGAGAGGTGTGATCGCTCGAGCAGAGGTGAACCC 879
Db 991 GSGATGAGCTAGGCGAGCCATTGTCCAGAAAAGTCCCTATATCATCAAGCCGAAATTAATCC 1050
QY 880 GTACCGGATGATCGTGTGTGCTGCTGTGTGCTGCGCTCTTCTCCGGTACCGTAT 939
Db 1051 CTACAGATGATATCGTTATCGGTTGCGTTGGTTTGGGTTTCTTCTCACTACCGAGT 1110
QY 940 CCTGCAACCCGTCGCGAGCCATCGGGCTGTGGCTGTCTCCATCATCTGCGAGATCTG 999
Db 1111 GATGCAATCCGGCGAAGATGCAATTTGATTTGGCTCATATCTGTAATCTGTGAATCTG 1170
QY 1000 GTTCGCCATCTCTCGATCTCGACCGAGTTCCCAAGTGGTTCCCAATCGACCGGAGAC 1059
Db 1171 GTTTGCGATGCTGGGATTTTGATCAAGTTCCCAAGTGGCTTCCAATCGAGAGAGAGAC 1230
QY 1060 GTACCTCGACCGCTCTCCCTCAGGTACGAGAGGAGGGAGCGCTGCTGCTGCTGCGG 1119
Db 1231 TTACCTGGACCGTTTGTCACTAAGTTTGAAGAGAGGTCAACCTCTCAGCTTGTCTCC 1290
QY 1120 GGTGGACCTGTTCTGTGAGCAGGTGGACCGCTCAAGGAGCCGCGTGGTGAACCGCAA 1179
Db 1291 AATCGACTCTTTGTGAGTACGTTGATCCCAAGAGGAACTCCCTTGGTCAACGGAA 1350
QY 1180 CACCGTGCTCTCATCTCCCGTAGACTACCGGTGGACAAGGTCTCTGCTACGTCTC 1239
Db 1351 CACTGTCTTTTCCATCTTTCTGTGGATTATCCGGTTGAGAGGTTCTCTGCTATGTTTC 1410
QY 1240 CGAGCAGGCGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1299
Db 1411 TGATGATGGTGTGCAATGCTTACGTTTGAAGCATGTCTGAAACATCTGAAATTTGAAA 1470
QY 1300 CAAGTGGGTCCCTTTCTGCAAGAGTTTCGGCATCGAGCCCCCGCGCGAGTTCTACTTT 1359
Db 1471 GAAATGGGTCTCTTTCAGCAAAAAGTTTAATATCGAGCTCGTCTCTGAGTGGTACTT 1530
QY 1360 CTGCTCAAGTGTGACTACCTCAAGGACAGGTGCGAGCCCACTTCTGTGAGGAGCGCG 1419

Db 1531 CCAACAGAAAGATAGACTACTCTGAAAGCAAGGTGCTGCTTCAATTTGTTAGGAGAGGAG 1590
QY 1420 CGCCATGAAGAGAGAGATGATGAGGAGTTCAAGGTCCGGATCAACCGCTGTTGGTCCCAAGGC 1479
Db 1591 GGGGATGAAGAGAGATACGAGGATTTCAAGGTAGGATCAATGCCCTTGTGTTGCAAAAGC 1650
QY 1480 CATGAAGGTGCGCGCAGAGGGGTGGATCATGAAGGACGCAACGCGTGGCCCGGGAAACAA 1539
Db 1651 CCAAAAGGTTCTGTAGGAAGGATGGACAATGCAAGATGGAAGCCCTCGCTGGAACAA 1710
QY 1540 CACCCGCGACCAACCCCGGCATGATCCAGSTGTTCTTGGGCCACAGCGGCGGCACACAC 1599
Db 1711 CGTACGCGATCATCTCTGGAATGATTCAGGTATTTCTTGGCCAAAGTGGCGGTGCTGATGT 1770
QY 1600 CGAGGCAACGAGCTGCCCGCTCGTGTGACGTCTCCGTGAGAACGCGCCGGGATTTCCA 1659
Db 1771 GGAAGGAATGATGTTGCTCGCTGCTGTTATGTCTCGAAGAAAGAGGCGAGGTATTA 1830
QY 1660 GCACCAAGAAAGCGCGCCCATGAACGCTCTGATTTCCGCTCTCCCGCTGCTGACCAA 1719
Db 1831 CCATCAAGAAAGGCTGGTGCATGAATGCACTGGTCCGCTCTCTGCTGTCTTATCAA 1890
QY 1720 CCGGCCATTCATGCTCAACTTTGGACTGTGATCACTACATCAACACAGCAAGGCCATCCG 1779
Db 1891 TGCCTGATACCTATTGAACTTTGGACTGTGATCACTACATCAACATAGCAAGGCCATAA 1950
QY 1780 GAGGCGCATGTCTTCTCATGGAACCTCAGGCTCGGCGGAGAGGTCTGCTACGTTCACTT 1839
Db 1951 AGAGGCTATGTTTTCATGATGATCTTTGTTGGGAGAAAGTGTGCTATGTACAGTT 2010
QY 1840 CCGCAGAGGTTGCAAGCGCATCGAGTGCACGACGATACGTAACGTAACAGAAACACGTCCT 1899
Db 2011 CCTCAGAGGTTTGTATGTTGATTAACAAAATGATCGATACGCTAAACAGAACTGCTCTT 2070
QY 1900 CTTGCAATCAACATGAAGGGCTGGAACGCTACCAACCTCCCAAGGAGCCCAAGAGGCC 2019
Db 2071 TTTTGACATCAACATGAAGGTTTGGACGTTATCAAGGACCCATTTATTTGGGTACTGG 2130
QY 1960 GTGCGTGTTCGCGCGCAGCGCTCTACGGCTACAACTCCCAAGGAGCCCAAGAGGCC 2019
Db 2131 ATGTGTTTTGACAGCGCAGGCTGATGATGTTATGATGCTCTTAAACGAGAGCCACC 2190
QY 2020 -----CAAGATGGTGACCTGCGACTGTGCTCCGCTGCTCGGCGCGCAAGAGCG 2067
Db 2191 ATCAAGAACTTGCAACTGCTGCGCCAAAGTGGTGCCTCTCTTGTGCTGCGCAGGAGCAA 2250
QY 2068 GAAACAGCCCAAGGACGGGCTGCGGAG----- 2095
Db 2251 GAATAAAAAGAGACTACAAAACCAAGACGGAGAGAAAGATTTATTTTCAAGAA 2310
QY 2096 -----GGCACCGCTGATATGGAGT 2115
Db 2311 AGCAGAAAACCATCTCTGCTATATGCTTTGGGTGAAATTTGATGAAGGTGCTCCAGGTGC 2370
QY 2116 AGATAGCGCAAGAGAGATGCTCATGTGCCACATGAATTTGAGAAAGCGTTGCGGCGAGTC 2175
Db 2371 TGATATCGAAGAGGCGGAAATCGTAAATCAACAGAAATAGAGAGAAATTTGGCGAGTC 2430
QY 2176 CCGCGGTTTCGTCAAGTGCAGCTGATGGAGGAAGGCGGCTCCCTCTCTGCTGAGGCC 2235
Db 2431 TTTCTGTTTTTGTGCGCATCAACACTTCTTGAGAACGAGGAGCCCTGAAGAGCGCAAGTCC 2490
QY 2236 CCGCGGCTCTCAAGGAGGCGCATCCATGTCTAGCTGCGGTACGAGAGCAAGACCGA 2295
Db 2491 AGCTTCTCTCTGAGGAAGCTATACATGTTATCAGCTGCGGTACGAGAGCAAGACCGA 2550
QY 2296 CTGCGGCTGAGCTGGGTGATCTAAGGCTGATCAAGGAGCAATCTCTGAGCGGGTT 2355
Db 2551 CTGGGAAAAGAGATGGCTGATTTTACGATCGATCAAGAGGATATCTTGACTGGATT 2610
QY 2356 CAAGATGCACTCGCGGCTGCGCTCCGTGCTACTGATGCGAGCGGCGGCTTCAA 2415

Db 2611 TAAGATGCACTGCCATGGCTGGCGTCTATTATTGTCATCCGAGCGGCTGTCAATCAA 2670
Qy 2416 GGGGTGGCGCGGATCAATCTATCGGACCGTCTCAACGAGGTGTCTGGGTGGCGCTGGG 2475
Db 2671 AGTTTCTGGCGCTCTGAACCTTTCCGACCGTCTTCCAGGTCCTTGGTGGCGCTTGG 2730
Qy 2476 GTCCGTGAGATCTTCTTCCAGCGGCGACAGCCCGCTGTGTACGCTACAGAACGGCAA 2535
Db 2731 GTCCGTGAAATTTCTTCCAGAACGACTGCCCACTTTGGTACGGATAC---GGCGGCGG 2787
Qy 2536 CCTCAAGTGGCTGGAGCGCTTCGCTCATCAACACACACCATCTACCCCTTTCACCTCGCT 2595
Db 2788 GCTAAATCTCTGAAGAGTTTCTTATATCAACTCCATCGTTATCCCTGGAGCTCAT 2847
Qy 2596 CCGCTGTCTGCTTACTGACACCTCCCGCGCTGTGCTCTCAACCGGCAAGTTTCATCAT 2655
Db 2848 TCCTCTCTGCTTACTGTACTTCTGCTGCTGCCATCTGCTGTCAAGGGAAGTTTATCAC 2907
Qy 2656 GCGGTGATTAGCAGTTTCGCGAGCTTCTTCTTATGCGCTTCTCATGCTCCATCTTCCG 2715
Db 2908 ACCAGAGTTTACCAGTTTGGCCAGTATCTGGTTTATGGCATTTTTCATCTGCATCTCCGT 2967
Qy 2716 GACGGGATCTCTGAGATGCGGTGGAGCGGGTGGAGCATCGAGAGTGGTGGAGAACGA 2775
Db 2968 GACCGGATCTCTGAATAGTGGAGTGGGTGGCCATCGACCATGTGTGGAGAACGA 3027
Qy 2776 GCAGTCTGGGTATCGCGCGGTGTGCGGCGATCTCTTTCGCGCTGTGAGAGGCTGTCT 2835
Db 3028 GCAGTCTGGGTATCGAGGCGTTTCGGCGCATCTGTTCGCGGTTCAGAGGCGCTGCT 3087
Qy 2836 CMAGTCTCGCGGGATCGACACCACTTACCGTCACTTCCAAAGGCCACCGCGGACGA 2895
Db 3088 GAAGGTGTTCGCGGCGATCGACAGAGCTTCCACCGTCAAGTTCGAGGCGCGGAGACGA 3147
Qy 2896 GGACGAGAGTTCCCGAGCTCTACGCTTCAAGTGGACACAGCTCTCTCATCCCGCCAC 2955
Db 3148 G-----GAGTTCTCGAGCTGTACAGTTCAAGTGGACACCTCTGTATACCCCGAC 3201
Qy 2956 CAGCTGCTCATCATTAACGTATCGCGGTGTGGCGGATCTTCCGACGCGCATCAACAA 3015
Db 3202 CACGCTCTCTCTGCTGAATTTATCGGGGTGTGGCGGATCTCGAACGCGATCAACAA 3261
Qy 3016 GGGTACAGTCTCGGGGCGCTTTCGCGAGCTTCTTTCGCTTCTGGGTCTATCGT 3075
Db 3262 CGGTGAGAGTGTGGGCGCGCTTTCGGAAGCTTCTTTCGCTTCTGGGTGATCGT 3321
Qy 3076 CCACCTCTACCGGTCTCTCAAGGGGCTCATGGGCGGCGAGAACAGGACGCCACCGTTGT 3135
Db 3322 CCACCTGTACCGGTCTCTCAAGGGTCTGGTGGGAGGAGAGAACAGGACGCCAGTCTGT 3381
Qy 3136 TGTCACTGTGTCATCTGTGCGCTCATCTTCTCCCTGTCTTGGGTGAGGATCGACCC 3195
Db 3382 CATGCTGTGGTCCATCTCTGTCGCTCGATCTTCTCTGCTCTGTGGGTCTCGGCTCGACCC 3441
Qy 3196 TTTTCATGTGAGGACCAAGGCGCGGAGCTCAGGACGTGTGGCATCAATTGC 3247
Db 3442 GTTCTCGCCCAAGACCAAGCGCGCTCTCTGAGAGTGTGGCTGTGACCTGC 3493

RESULT 10

AAS16458

ID AAS16458 standard; cDNA; 3799 BP.

XX

XX

XX 14-FEB-2002 (first entry)

XX

XX Corn cDNA encoding cellulose synthase Cqrae19/cesa-19.

XX

XX Corn; ss; cellulose synthase; Cdpgs45; cesa-3; Cqrae19; cesa-9;

XX stalk quality; improved stand; silage; pericarp; kernel hardening;

XX handling ability; transgenic plant.

XX

Qs Zea mays.

XX Key Location/Qualifiers
FT CDS 238..3799
FT /*tag= a
FT /product= "Cellulose synthase"

PN W0200179516-A2.

XX 25-OCT-2001.

XX 12-APR-2001; 2001WO-US011951.

XX 14-APR-2000; 2000US-00550483.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG;

XX WPI; 2002-041338/05.

XX P-PSDB; AAU10496.

XX New cellulose synthase polypeptides and polynucleotides, useful in improving stalk quality or silage, and in increasing concentration of cellulose in the pericarp, hardening the kernel for improved handling ability.

XX Claim 4; Page 80-85; 88pp; English.

XX The invention relates to isolated nucleic acids encoding two cellulose synthase proteins from corn, Cdpgs45 (cesa-3) and Cqrae19 (cesa-9). Also disclosed are a recombinant expression cassette comprising the polynucleotide (operably linked to a promoter) a host cell comprising the recombinant expression cassette and a transgenic plant comprising the recombinant expression cassette. The nucleic acid is useful in the improvement of stalk quality for improved stand or silage, and in the increased concentration of cellulose in the pericarp, hardening the kernel, and thus improving its handling ability. The nucleic acids may also be used as probes or amplification primers in the detection, quantification or isolation of gene transcripts, as probes in detecting deficiencies in the level of mRNA, for detecting gene mutations or allelic variants, for monitoring up regulation of expression or changes in enzyme activity in screening assays, for site directed mutagenesis, and in sense or antisense suppression of one or more genes in a host cell, tissue or plant. The polypeptides may be used in assays for enzyme agonists or antagonists, and as immunogen or antigen to obtain antibodies specifically immunoreactive with the protein. The present sequence encodes a corn cellulose synthase of the invention

SQ Sequence 3799 BP; 925 A; 931 C; 995 G; 944 T; 0 U; 4 Other;

Query Match 38.4%; Score 1323.4; DB 6; Length 3799;
Best Local Similarity 65.3%; Pred. No. 1.1e-209;
Matches 2132; Conservative 0; Mismatches 986; Indels 147; Gaps 7;

Qy 133 GAACGAGTGGTGTCTGATCCGGGCGCACGAGGACCCCAAGCCGCTGCGGGCGGTGACCGG 192

Db 225 GAGCTCGCTGCCATCGAGGGCGACGCGGACGCGTGAAGTTCGGGGAGGCGCGGGGAGG 284

Qy 193 GCAGGTGTGCGAGATATCGGCGACGAGGTTCGGGCTTCAGGTGAGCGGACCTCTTCGT 252

Db 285 GCAGGTGTGCGAGATATCGGCGGATCGCGTGGGCACTACCGCGGAGGAGACGCTTTCAC 344

Qy 253 CGCTCTGCAACGAGTGGCGCTTCCCGGTGTGCCCGGCTCTGTCAGAGTACGAGCGCGGGA 312

Db 345 CGCTCTGCGAGCTCTGGGGTTCGCGGTGTGCCCGCTCTGTCAGTACGAGGCGCAGGA 404

Qy 313 GGCACGCGAAGTGTCCCGCAGTGCAGAGCGCGCTCAAGCGCTCAAGGGGAGCCGAG 372

Db 405 CGGCACACAGCGTGGCCCCCAGTGCAGGCAAAACAAGTACAAGCGCCACAAGGGGAGTCCAGC 464

Qy 373 GGTTCGGGGACGATGACGAGGAGGACATCGACGACCTCGAGACGAGTTCACATC-- 430

Db 465 GATCCGAGGGGAGGAAGGAGACGATACTGATGCCGATGATGCTAGCGACTTCAACTACCC 524
QY 431 -----GACGACGAGAAATCAGCAGAGGCGAGCTGGAGGGCAACATCAGACAACGCGAGAT 483
Db 525 TGCATCTGGCAATGACGACCAAGAGCAGAAGATTGCTGACAGGATCGCGAGCTGGCGCAT 584
QY 484 CA-----CCGAGGGGATGCTGCA CGCGCAGGATGAGCTACGCGGAGGGGCCCCG--- 530
Db 585 GAATGCTGGGGGAGCGGGGATGTTGGCGGCCCAAGATGATGACAGTGGTGAATCGGGCT 644
QY 531 -----ACGACGGCGACGGCAACAACACCCCGCAGATCCCGCCCATCATCACCGGCTC 582
Db 645 TACCAAGTAGCAGAGTGGTGAATCCCTCGGGGATACATCCCGTCAGTCACTAACAGCCA 704
QY 583 CCCTCCGTCGCCGCTGAGCGGTGAGTTTCCGATTCACCAACGGGTATGGCCACCGCGAGGT 642
Db 705 GATTTCGGGAGAAATCCCTGGTGTCTCCCTGACCATCATATGATGTCCTACTGGGAA 764
QY 643 CTGCTCTTCCCTGCA CAAGCGCATCCATCCGTACCTGTGTCTGAGCCAGGGAGTGCCAA 702
Db 765 CATTCGCGAGCGCGCCCATTTCCCTATATGAATCATTCATCAATCCCGTCGAGGGAATT 824
QY 703 GTGGGAGGAGAAAGAAAGTAGCTGGAAGGAGAGGATGGACGACTGGAAGTCCAGCA 762
Db 825 CTCTGTAGCGTTGGGAATGTTGCCTGGAAAGAGAGGGTTGATGGCTGGAAAATGAAGCA 884
QY 763 GGCA----- 767
Db 885 GGAACAGGGNA CAATTCCTCCATGACGAATGGCAACAGCATTGCTCCCTGAGGGCGGGG 944
QY 768 --TCCTGGCGGGCGCGGATCCGGAAGACATGSA CGCGACGTGGCATGCAACGCA 825
Db 945 TGTGTGTGATATTGATGATCACTCAACTGATTACACATGGAAGATGCCCTTATTAAACGATGA 1004
QY 826 GGGAGGCGACCGCTGTGAGGAAGGTGTCGATCGCGTCGAGCAAGGTGAACCCGTACCG 885
Db 1005 AACTCCCGACGCTTATAGGAAAGTTCCACTTCTTCTCCAGGATAAATCCATACAG 1064
QY 886 GATGGTGATCGTGGTGTCTCGTGTGTGCTCGCTTCTTCTCCGCTACCGTATCTGCA 945
Db 1065 GATGGTCATGTGCTAGGATTGATTGTTTAAGCATCTTCTTGACATCTCCGATCACAA 1124
QY 946 CCCCGTCCGGACGCCATCGGGCTGTGGTCTGCTCCATCATCTCGAGATCTGGTTCG 1005
Db 1125 TCCTGTGCGTAATGTCATACCACTGTGGCTTCTATCTGTATATGTGAGATCTGCTTGC 1184
QY 1006 CATCTCTGTGATCTCGACCACTTCCCAAGTGGTTCGCCATCGACGGGAGACGTACCT 1065
Db 1185 TCCTTCTGGATATTGGATCAGTTTCCAAAGTGGTTTCCAAATCAA CCGGAGACTTACCT 1244
QY 1066 CGACCGCTCTCCTCAGGTACGAGAGGGAAGGGAGCGCTCGCTGTGTCGGCGGTGA 1125
Db 1245 TGATAGACTCGCATTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGTGTCTTGA 1304
QY 1126 CTTGTTGTGAGCACCGTGGACCCGCTCAAAGAGCCCGCTGGTGAACGCCCAACCCGT 1185
Db 1305 CATTTTGTGAGTGTGACCCCAATGAAGGAGCCCTCCTTGTCACTGCCAATACCGT 1364
QY 1186 GCTCTCCATCTCGCGGTAGACTACCCGTGGACAGAGTCTCTGCTAGCTCCGACGA 1245
Db 1365 GCTATCCATCTCTCGCTGTGGACTATCCTGTGGAATAAGGTCTCTTGTATGTATCTGATGA 1424
QY 1246 CGCGGCTCGATCTGACGTTTCGAGTCTGCTCGGAGAGCGCGAGTTGCGCGCAAGTG 1305
Db 1425 TGAGCTGCTATGCTGACATTTGATGCACTAGCTGAGACTTTCAGAGTTTGCTAGAAATG 1484
QY 1306 GGTGCCCTTCTCAAGAAGTTCGGCATTCGAGCCCGCGCCCGAGTTCTACTTCTCGCT 1365
Db 1485 GGTGCCATTTGTAAGAAGTACAACATTGAACCTTAGAGCTCCTGAATGGTACTTCTCCCA 1544
QY 1366 CAAGGTGCACTACCTCAGGACAAAGGTGAGGCCCACTTCTGTCGAGAGCGCGCGCAT 1425
Db 1545 GAAAAATTGATTACTTGAAGGACAAAGTGCA CCCTTCAATTTGTTAAAGAACCGCGCGGCCAT 1604

QY 1426 GAAGGAGAGATATGAGGAGTTCAAGGTCGGATCAACGCGCTGGTGGCCAGGCCATGAA 1485
Db 1605 GAAGGAGAAATATGAAGAATTCAAATTAGGCTAAATGGCCCTTGTGTGAAGCAAAA 1664
QY 1486 GTTCCGGCAGAGGGGTGATCATGAAGGACGGCAGCGCTGGCCGGGAAACAACACCCG 1545
Db 1665 AGTCCCTGAGGAAGGATGGATCATGCAAGATGSCACCATGCGCCAGGAACAATACCA 1724
QY 1546 CGACACACCCCGCATGATCCAGGTGTTCTGTGGCCACACAGCGCGCGCACACCGAGG 1605
Db 1725 GGACCATCTCTGAATGATTCAAGTTTCTTGGTCAAGTGGTCTTCTGATCTAGAGG 1784
QY 1606 CAAAGAGCTGCCCGCTCGTGTAGTCTCTCCGTGAGAAGCCCGGGATTCAGCACCA 1665
Db 1785 TAATGAGCTACCCGTTGGTCTATGTTCTCGTGAATAAAGCTCCTGGATTCCAGCATCA 1844
QY 1666 CAAAGAGCGCGCGCATGAACGCTCTGATTCGCTCTCCGCGCTGTGACCAACCGCGC 1725
Db 1845 CAAAGAAAGCTGGTGCATGAATGCTTGTCCGCGCTCTCAGCTGTGCTTACCAATGACA 1904
QY 1726 ATTCATGCTCAACTTGGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGGC 1785
Db 1905 ATACATGTTGAATCTTGAATGATCACTACATCAACACAGTAAGGCTCTCAGGNAAGC 1964
QY 1786 CATGTGCTTCTCATGAGCCCTCAGGTCCGCGGAAAGGTCTGCTAGTTCAGTTCCCGCA 1845
Db 1965 TATGTGCTTCTTATGATCCCTAAGTAAAGGAGTGTCTGCTATGTTTCAAGTTTCCCCA 2024
QY 1846 GAGGTTTCGACGGCATCGAGTCGACGACGATACGCTAAACGTAACGAGAACACCGTCTTTCGA 1905
Db 2025 GAGGTTTCGATGTTGATAGGAATGATCGATATGCCAAACAGGAACACCGGTGTTTTCGA 2084
QY 1906 CATCAACATGAAGGGCTGGACGGCATCCAAGCCCGGTGTACGTCGGGACAGGGTGCCT 1965
Db 2085 TATTAACTTGAAGGCTTGTGATGAGGATCCAGGACAGTATATGTTGGGACCTGGCTGTGT 2144
QY 1966 GTTCCGGCGCCAGGGGCTCTACGGCTCAACCCCTCCCAAGGACCAAGGCGCCCAAGAT 2025
Db 2145 TTTCAACAGAACAGCTCTATATGTTATGAGCCGCCCAATTAAGCAAAAAGGGTGGTTT 2204
QY 2026 GGTGACTCGGACTGTGTCGCCGCTGTTCGGCCGCAAGAGCGGAAACACGCCAAGGAC-- 2083
Db 2205 CTTGTATCACTATGTGTGGTGGCAGGAAGGAAAGCAAAATCAAAGAGGGGTCAGACAA 2264
QY 2084 -----GGGCTGCC 2091
Db 2265 GAAAAAGTCAAGAAAGCATGTGACAGTTCGTGCCAGTATTCAATCTTGAAGATATAGA 2324
QY 2092 GGAGGGCAACCGCTGATATGGGAGTAGATAGCAGCAAGGAGATGCTCATGTCCCACATGAA 2151
Db 2325 GGAGGGAGTTGAAGCGCTGGATTGATGATGAGAAATCACCTTCTTATGCTCAAAATGAG 2384
QY 2152 CTTTCGAGAAAGCGGTTGCGGACGTCCGCGCGCTTGTGCTGACGCTGATGAGGAAAGG 2211
Db 2385 CTTGGAGAGAGATTGGGCAATCTGCAGCTTTTGTGTGCTCCACTCTGTAAGAAATGG 2444
QY 2212 CGCGCTCCCTCTCTGTCGAGCCCGCGCGCTCTCAAGGAGGCCATCCATGTCAATCAG 2271
Db 2445 TGGTGTCTCTCAGTCTGCGACTCCAGAAATCTCTTCTTGAAGAAGCTATCAATGTCAATAG 2504
QY 2272 CTGCGGCTACGAGGACAAAGCCGACTGGGGGCTGGAGCTGGGGTGGATCTACGGGTGAT 2331
Db 2505 TTTGGCTACGAGGACAAAGATTGAATGGGAACTGAGATTGGGTGGATCTATGTTCTGT 2564
QY 2332 CACGAGGAGCATCTGACCGGGGTTCAAGATGACCTGCCCGGGGTGGCGCTCCGTGTACTG 2391
Db 2565 GACGGAAGATATTCTCACTGGGTTCAAGATGACGCA CAGAGGCTGGCGGTCTACTCTG 2624
QY 2392 CATGCCGAGCGGGCGGCTCAAGGGGTCGCGCGGATCAATCTACTCGACCGCTCTCAA 2451
Db 2625 CATGCCTAAGCGGCGCGGCTTCAAGGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAA 2684


```
Qy 2452 CCAGGTGCTCGGTGGGGCTGGGGTCCGTGCGATCTTCTTACGCGGCACAGCCCGCT 2511
Db 2685 CCAGGTGCTCGGTGGGGCTCGGTTCAGTGAATCTTTTACGCGGCATTTGCCCGCT 2744
Qy 2512 GCTGTACGGCTACAAGAACGCAACCTCAAGTGGCTGGAGCGCTTCCGCTACATCAACAC 2571
Db 2745 ATGTTACGGGTAC---GGAGGACCGCTGAGTTCTTGGAGAGATTCGCTACATCAACAC 2801
Qy 2572 CACCATCTACCCCTTCACTCGCTCCGCTGCTCGCTTACGCAACCTTCCCGCGGTCTG 2631
Db 2802 CACCATCTACCCGCTCAGTCCGCTCCGCTCCTCATTTACTGTATCTGCTGCGCATCTG 2861
Qy 2632 CTTCTCTACCGGCAAGTTCTATCATGCGCTCGATTAGCACGTTGCGCAACCTCTTCTCAT 2691
Db 2862 CTTGCTCAGCGGGAAGTTCTATCATCCAGAGATCAGCAACTTCGCTAGTATCTGTTTCAT 2921
Qy 2692 CGCCCTCTTATGTCCTATCTTCGCGAGCGGTCATCTGAGATGCGGTGGAGCGGGTGAG 2751
Db 2922 CTCTCTCTTCTGATCTTCTGCGACGCGGTATCTTGAGATGAGTGGAGCGGGTGGG 2981
Qy 2752 CATCGAGAGTGTGGAGGAACGACAGTTCTGGGTTCATCGCGCGGTGTCGCGCATCT 2811
Db 2982 CATCGAGAGTGTGGAGGAACGACAGTTCTGGGTTCATCGGAGGTCATCTCGGCCACCT 3041
Qy 2812 CTTGCGCGTGTGAGGCGCTCTCAAGTCTCTCGCGGATCGATCGCAACCTTACCGT 2871
Db 3042 CTTGCGCGTCTTCCAGGCGCTCTCAAGTCTCTCGCGGATCGATCGCAACCTTACCGT 3101
Qy 2872 CACTCTCAAGGCGCACCGCGAGGAGGACGAGTTCGCGAGTCTTACGCTTCAAGTG 2931
Db 3102 CACTCTCAAGGCGCTG---GATGAGACGGGACTTCGCGAGTGTACATGTTCAAGTG 3158
Qy 2932 GACCACGCTCTCTATCCCGCCACCGCTGCTCATATTAACGTCTATCGCGCTGCGGCG 2991
Db 3159 GACGACACTTCTGATCCCGCCACCAACCATCTGATCATCAACCTGCTGCGGTGTTGTC 3218
Qy 2992 CGGCATCTCGACCGCATCAACACGGGTACCACTCTGGGGCGCTCTTTCGCAAGCT 3051
Db 3219 CGGCATCTCTTACCGCATCAACACGGGTACCACTCTGGGGTCCGCTCTTTCGCAAGCT 3278
Qy 3052 CTTCTTCGCTTGGGTGATCGTCCACCTCTACCGTCTCTCAAGGGCTCATGGGGCG 3111
Db 3279 CTTCTTCGCTTGGGTGATCGTTCACCTGTCACCGTCTCTCAAGGGTCTCATGGGTGCG 3338
Qy 3112 CCAGAACAGGACGCCCGCTGTTGTTGTCATCTGTTCCATTTCTGCTGCGCTCCATCTTCTC 3171
Db 3339 GCAGAACGACCGCACCGGACCATCGTGGTGTCTGCGGATCTCTGCGGTGATCTTCTC 3398
Qy 3172 CTTGCTCTGGGTAGGATCGACCTTTTCATGTCAGGACCAAGGGCGCGGACGTGAGCA 3231
Db 3399 CTTGCTGTTGGTTTGGCATCGATCGTTCACCAACCGGTCTACTGGCGCGGATCTCGAAC 3458
Qy 3232 GTGTGGCATCAATTGCTGAGCTGT 3256
Db 3459 GTGTGGCATCAACTGCTAGGAGGT 3483
```

RESULT 11

AAZ58263

ID AAZ58263 standard; cDNA; 3776 BP.

XX

AC AAZ58263;

XX

DT 08-MAY-2000 (first entry)

XX

XX Corn cellulose synthase cDNA.

DE

XX Corn; maize; cellulose synthase; transgenic plant; ss.

KW

XX

OS Zea mays.

XX

FH Key Location/Qualifiers

FT CDS 2..3448

FT
XX
PN
XX
XX
PD
XX
PF
XX
PR
XX
XX
PA
XX
PI
PI
XX
DR
DR
XX
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ

WO200004166-A2.

27-JAN-2000.

13-JUL-1999; 99WO-US015871.

14-JUL-1998; 98US-0092844P.

(DUPO) DU PONT DE NEMOURS & CO E I.

Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;

Rafaleki JA, Thorpe CJ;

WPI; 2000-182431/16.

P-PSDB; AAY58832.

Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as probes for isolating cDNAs and genes encoding homologous proteins, for producing transgenic plants.

Claim 9; Page 37-38; 93pp; English.

The present sequence is that of a contig of cDNA clones and encodes a portion (see AAY58832) of corn cellulose synthase (CS). The cDNA clones were isolated from corn cob, developing tassel, seedling, shoot culture and leaf tissue cDNA libraries on the basis of homology to Arabidopsis and cotton CS sequences. The invention relates to isolated nucleic acid fragments encoding plant CS and to CS polypeptides. It also relates to the construction of a chimeric gene encoding all or a portion of the CS, in sense or antisense orientation, where expression of the gene results in altered levels of the CS in transformed host cells. The host cells can be used to screen compounds for their ability to inhibit CS activity. CS nucleic acids are also useful for producing transgenic plants having altered levels of CS, and hence altered levels of fibre. CS may also serve as a target for the development of novel herbicides

Sequence 3776 BP; 931 A; 906 C; 999 G; 940 T; 0 U; 0 Other;

Query Match

36.8%; Score 1268; DB 3; Length 3776;

Best Local Similarity 64.0%; Pred. No. 1.7e-200;

Matches 2097; Conservative 0; Mismatches 1055; Indels 126; Gaps 7;

Qy 87 GCCTGATGAGGCGCAGCGCGGCTGCTGCGCGCTCGCAACCGGAAAGAGTGTGC 146

Db 174 GGGGGATGAGGCGCAGCGCGGCTGCTGCGCGCTCGCAACCGGAAAGAGTGTGC 233

Qy 147 TGATCCGGGGCCACGAGGA-----CCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGT 200

Db 234 TCATCCGCGCGATGCGAGCCAGGGCGCGAAGCCCATGGACCAAGCGGAAACGGCAGGTGT 293

Qy 201 GCGAGATATGCGCGGACGAGGTGCGGCTCAGGTGGAACCGGACCTTCTGTCGCTGCA 260

Db 294 GCCAGATTTGCGGCGACGAGCTGCGGGCGCAACCCCCACGGGGAGCGGCTTCTGCGCTTCA 353

Qy 261 ACAGTGGCGCTTCCCGCTGTCGCGCCCTGTCTACGAGTACGAGCGCGGGAGGGCACGCG 320

Db 354 ACAGTGGCGCTTCCCGCTGTCGCGGACTGTCTACGAGTACGAGCGCGGGAGGCAAGC 413

Qy 321 AGAACTGCCCGCCAGTGCAGAGACCGGTACAAGCGCTTCAAGGGAGCGCGGTTGCGCG 380

Db 414 AGAACTGCCCGCCAGTGCAGAGACCGGTACAAGCGCTTCAAGGGAGCGCGGTTGCGCG 473

Qy 381 GGGACCATGACGAGGAGGACATCGACACTTGGACGACGAGTTCACATCGACGACGAGA 440

Db 474 GGGACGAGGAGGAGGACGCGCTCGACGACTGGAGAACGAGTTCAACTGGAGGCAAGC 533

Qy 441 ATCAGCAGAGGAGCTGGAGGGCAACATGACAGAACGAGCCAGAT----- 483

Db 534 ACAGCTCCCGTACCTTCCGCGAGTCCATGCTCCAGCCCAATGAGCTACGGCGCGGCG 593

Db	462	GATGACATCCGCGCGGACCGCCCTTGTGCCCTCGTTCGGTGGCGGGGGAAG	521	1542	TCAGAGAAAAACGGCCAGGCTACAAACCATCAAAAGAAAGGCTGGTCTATGAATGCATTG	1601
Qy	614	ATTACCAACGGGTATGGCCACGGCGAGGTCTCGTCTTCCTCGTCAACGAGCCATCCATCCG	673	1694	ATTTCGGCTCTCGCGCTGCTGACCAACAGCGCCATTATCATGCTCAACTTGGACTGTGATCAC	1753
Db	522	AGGATTCACCCCTCTCCCGTAGCGGATCCCAACCTTCCTCTGTGCAACCGAGGTCTATGGAC	581	1602	GTCCGAGTCTCTGCTGTACTAACTAATGTCTCTTATTTGCTGAACTTGGATTGTGATCAC	1661
Qy	674	TACCTGTGTCTGAGCGAGGAGTGCCAAAGTGGGACGAGAAAGAAAGTGAAGTGAAG	733	1754	TATATCAACAAACAGCAAGGCCATTCGGGAGGCCATGTCTCTCTCATATGAAACCTCAGGTC	1813
Db	582	CTTTCAAGGATCTCGCGCATATGGCTACGGGAGCGTAGCATGGAAGGAGAGATGGAG	641	1662	TATATCAATAATAGTAGTAAGGCTATAAAGGAAGCAATGTGTCTTATGATGATCCTTTGCTT	1721
Qy	734	GAGGGATGACAGTCTGGAATCCAGCAGGGGATCCTCGCGCGCGCGCGATCCGGA	793	1814	GGCCGAAGGTCTGCTACGTTTACGTTCCGAGAGGTTTCGACGGCATTCAGCTGCAGAC	1873
Db	642	AGCTGGAAGCAGAGCAGGAGAGATGACACAGCAGAGAACCATGCGCGCGCGGATGAT	701	1722	GGAAAGAAAGTTTGTCTATGTGAGTTTCTCAAGATTTTGAATGGGATTCATGCCCATGAT	1781
Qy	794	GACATGACCCCGACGTGGCACTGAACAGCAGCGAGGAGCGCGTGTTCGAGGAAGGTG	853	1874	CGATAGCTTAAACAGGAACACCGTCTCTTCGACATCAACATGAAGGGCTGGAACGCATC	1933
Db	702	GGTGATGATCAGATCTACCACTAATGGATGAAGCTAGACAGCAATGTTCGAAGAATC	761	1782	CGATATGCTTAAACAGAAATGTTGTCTTTTTCGATATCAACATGAAGGTTTGGATGGTATC	1841
Qy	854	TCGATCCGCTCGAGCAAGGTGAACCCGTACCGATGGTGATCGTGGTGCTCTCGTTGTG	913	1934	CAAGCCCGGTGTACGTTCGGGACAGGGTGTGTTCCGGCGCCAGGCGCTCTACGCTAC	1993
Db	762	CCGCTTCTTCAAGCCAAATCAACCCCTATAGGATGATTAATAATTCGGCTAGTGGTT	821	1842	CAGGCGCCAAATTTATGTGGGTACTGATGTGCTTCAGAAAGCAGGCATTATATGGCTAC	1901
Qy	914	CTCGCTTCTCTCCGTTACGTTACCTGTGACACCCGTCCTCCGACCGCATCGGGCTGTGG	973	1994	AACCTCCCAA-----	2004
Db	822	TTGTGTTTCTTCTCCACTACCGAGTGAATCATCCGCTGCTGATGCAATTTGCTTTATGG	881	1902	GATGCTCCCAAAACAAAGAACCAACATCAAGAACTTGCAACTGTGTGGCCAAAGTGGTGC	1961
Qy	974	CTCGTCTCCCATCATCTGCGAGATCTGTTTCGGCATCTCTTGGATCTCTCGACCAAGTTCCCC	1033	2005	-----GGGACCCCAAGAGCCCAAGATGGTGTGACCTGTGCTGCTGCTGCCCG	2047
Db	882	CTCATATCTGTGATCTGTGAATTTGGTTTGGCCATGTCTTGGATCTTTGACCAAGTTTCCA	941	1962	ATTGCTGTGTGTTTGGTAAACAGGAAGACCAAGAAAGAACCAAGACCTCTAAACCT	2021
Qy	1034	AGTGTGTTCCCATCGACCGCAGAGATCACTTCGACCGCTCTCCCTCAGSTACGAGAGG	1093	2048	TGCTTGGCGCGCAAGAAG-----CGGAAACACGCCAAAGAACGGGTGCGCGGAGGC	2098
Db	942	AAAGTGGTTTCTATCGAGAGGGAACCTATCTTTGACCGGCTGAGTTTAAAGTTTGACAAG	1001	2022	AAATTTGAGAAGATAAAGAACTTTTAAAGAAAGAAAGAAATCAAGCCCCTGCAATGCT	2081
Qy	1094	GAAGGGAGCGGTGCTGTGTCGGCGGTGGAACCTGTTTCTGTGAGCAGGTGGACCGCTC	1153	2099	ACGCTGATAT-----GGGAGTAGTAGCGACAAGAGATGCTCATGTCTC	2143
Db	1002	GAAGGGGATCCTTCTCAACTCGCCCCCTGTGTTTCTTGTCTAGTACGGTGTGATCCCTTG	1061	2082	CTTGGTGAATTTGATGAAGCCGCTCCAGGAGTGAATAATGNAAGGCTAGTATGTAAT	2141
Qy	1154	AAGGAGCGCGTGTGTGACGCGCAACACGCTGCTCTCCATCTCGCGGTAGACTACCCC	1213	2144	CACATGAACCTTCGAGAAGCGGTTTCGGGAGTCCGGCGCTTCGTACGTCGAGCTGATG	2203
Db	1062	AAGAACTTCATTTGGTCACTGTATATCTGTCTATCTATCTCTTTCGGTGGATATCCA	1121	2142	CAACAGAGTTGGAAAGAAATTTGGCAGTCTTCAGTTTTTGTGCAATCCACATCTT	2201
Qy	1214	GTGCAAGAGTCTCTGCTACGTCTCGACGACGGCGGTGATGCTGACGTTTCGAGTGC	1273	2204	GAGGAAGCGCGGTCTCTCTTGTGAGCCCCCGCGCTCTCTCAAGGAGGCCATCCAT	2263
Db	1122	GTTGATAAGGTTTCATGCTACGTTTCTGATGATGTGTGCTGCCATGTGTGAAGCA	1181	2202	GAGAAATGGTGAACCCCTGAAGAGTGCAGTCCAGCTTCTCTTCTGAAGGAAGCTATACAT	2261
Qy	1274	CTGTCGAGACGGCCGAGTTTCGCGCAAGTGGTGGCTTCTGTGCAAGAAAGTTTCGGCATC	1333	2264	GTCAATCAGTCTGCGGTACGAGGACAAAGACCGACTGGGCTCGAGCTGGGGTGGATCTAC	2323
Db	1182	TTGTCTGAAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAAAAGATATAGCCTT	1241	2262	GTCAATCAGTTGTGGATATGAAGACAAACAGGCTGGGGAAAAGATATTGGTTGGATTTAT	2321
Qy	1334	GAGCCCGCGCCCGGAGTTCTACTTCTCGCTCAAGTTCGACTACCTCAAGGACAAAGGTG	1393	2324	GGTTCGATCAACGAGGACATCCTGACGGGTTTCAAGATGCACTGCGCGGGTGGCGCTCC	2383
Db	1242	GAGCTCTGCTCAAGTGTGTTCTTCCAAAGAGATAGACTACCTGAAAGCAAGGTG	1301	2322	GAATCAGTCAAGAGATATTTCTTACTGGGTTTAAGTGCATGCGCATGGTTGGCGGTCA	2381
Qy	1394	CAGCCACCTTCGTGACGAGCGCGCCGATGAAGAGAGAGTATGAGGAGTTCAAGGTC	1453	2384	GTGTACTGCATCCGAAGCGCGGTTCGAAGGGGTTCGGCGCGGATCAATCTATCGGAC	2443
Db	1302	CGCCCAAACTTTGTTAGAGAACCGAGAGCAATGAAGAGAGATATGAGGAATTTCAAGGTC	1361	2382	ATTACTGCATACCTTAAACGGCGCGCTTCAAAGGTTCCGACCTCTCAATCTTCCGAT	2441
Qy	1454	CGSATCAACCGCTGTGGCCAAAGGCATGAAGGTGCGCGCAGAGGGGTGGATCATGAAG	1513	2444	CGTCTCAACAGGTGCTCGGTGGGGCTGGGGTCCGTGAGATCTTCTTCAGCCGGCAC	2503
Db	1362	AGAATCAATGCTTGGTTGTAAAGCCCAAAAGGTTCTCTGAGGAAGGATGGAACAATGCAG	1421	2442	CGTTTTACACAGGTTCTTCGGTGGGCTCTTGGTTCAATTTGAATTTTGTTCAGCAACCCAC	2501
Qy	1514	GACGGACCGGTGGCCCGGGAACAACACCGCGACCCACCGGCATGATCCAGGTGTTTC	1573	2504	AGCCCCCTGCTGTACGGCTACAAGAACCGCAACCTCAAGTGGCTGGAGCGCTTCGGCTAC	2563
Db	1422	GATGGAACTCCATGGCCCGGAATATATGTCCGTGATCATCTCGAATGATTCAGGTTTTC	1481	2502	TGCCCCCTCTGTTATGGGTATGGTGTGGA---CTAAGTTCTTGGAAAGGTTTTCGTAC	2558
Qy	1574	CTGGGCACAGCGCGCCACGACACCGAGGGCAACGAGCTGCCCGCTCTGTGTAGTCTC	1633	2564	ATCAACAACCAATCTACCCCTTCACTCGCTCCGCTGCTCGCTTACTGCAACCTTCCC	2623
Db	1482	CTTGGTCAAAAGTGGTGGCCATGATGTGAAGGAAATGAGCTGCTCGATTGGTTTATGTT	1541	2559	ATTAACTCCATGTTATACCTTTGGACATCTATCCGCTCTTGGCCCTTATTGCAATTTGCC	2618
Qy	1634	TCCGCTGAGAGCGCCCGGATTCAGCACCAAGAAAGCCCGCGCATGAAGCTCTG	1693	2624	GGGCTGTGCTCTCACCGCAAGTTTCATCATGCGCTGATGAGAGGTTTCGCCAGCTTC	2683
				2619	GCCATCTGCTGTGCTGACAGGGAAATTTATACGCCAGAGCTTTAAACAATGTTGCCACCTC	2678

Qy 2684 TTCTTCATCGCCCTTTCATGTCCATCTTCGGACGGCATCTCGAGATCGGTGGAGC 2743
 Db |||||
 Qy 2679 TGGTTCATGTCACTTTTCATCTGCATTTTGTCTACGACATCTCGAAATGAGATGGAGT 2738
 Db |||||
 Qy 2744 GGGGTGAGCATCGAGGAGTGTGGAGGAACGAGCAGTTCTGGGTTCATCGCGGTGTC 2803
 Db |||||
 Qy 2739 GGTGTAGCATCGATGACTGTGGAGAAACGAGCAGTTTGGGTCAITGGAGGGGTGTCT 2798
 Db |||||
 Qy 2804 GCGCATCTCTTCGCGCTCGTCAGGCGCTCTCAAGGTCTCTCGCGGATCGACACCAAC 2863
 Db |||||
 Qy 2799 TCACATCTCTTGTGTGTTCCAGGACTCTCTCAAGGTCTAGCTGTGTAGACACGAGC 2858
 Db |||||
 Qy 2864 TTCACCTGACCTCAAGGCCACCGGACGAGACGAGTTCGCCGAGCTCTACGCC 2923
 Db |||||
 Qy 2859 TTCACCTGTGACATCCAAAG-----GGCGGAGACGACGAGGAGTTCTCAGAGCTGTACACA 2912
 Db |||||
 Qy 2924 TTCAGTGGACACGCTCTCATCCCGCCACCGCTGCTCATCATTTAAGTGTATCGGC 2983
 Db |||||
 Qy 2913 TTCAAATGGACGACCTTCTGTATCTCGACAAACCTTCTCTACTGAACTTCAATGGA 2972
 Db |||||
 Qy 2984 GTCGTGGCGGCATCTCCGACGCCATCAACACGGGTACCAAGTCTCTGGGGCCCTCTTC 3043
 Db |||||
 Qy 2973 GTGTAGCTGCACTCTCAATGCGATCAACACGGATATGAATCATGGGGCCCTGTTC 3032
 Db |||||
 Qy 3044 GGCAGCTCTTCTGCGCTTCTGGGTCAATGTCACCTCTACCGTTCCTCAAGGGGCTC 3103
 Db |||||
 Qy 3033 GGGAGCTCTTCTTGTGATTTTGGGTGATCTGCTCCATCTTTACCGGTCTCTCAAGGGTCTG 3092
 Db |||||
 Qy 3104 ATGGGGCGCCAGAACAGACGCCACCGTGTGTCATCTGTCCTGTCCTGTCGCTCC 3163
 Db |||||
 Qy 3093 GTTGGAGGCGAAGACAGACGCCAACGATGTGTCATGTCGTGTCCTCTCTGCTTCG 3152
 Db |||||
 Qy 3164 ATCTTCTCCCTGCTCTGGGTGAGGATCGACCTTTTCATCGTCAGGACCAAGGGCCCGGAC 3223
 Db |||||
 Qy 3153 ATCTTCTCGCTGCTTGGGTGCGGATCGACCGTCTCTTGGAGGATGATGGTCCCTG 3212
 Db |||||
 Qy 3224 GTCAGGAGTGTGGCATCAATGC 3247
 Db |||||
 Qy 3213 TTGGAGGAGTGTGCTGTGATTGC 3236
 Db |||||

RESULT 13

ADA69779
 ID ADA69779 standard; DNA; 3264 BP.

XX AC ADA69779;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3102.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice; gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYNGEN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

XX PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

PS Claim 6; SEQ ID NO 3102; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

SQ Sequence 3264 BP; 784 A; 751 C; 852 G; 877 T; 0 U; 0 Other;

Query Match 35.2%; Score 1213.4; DB 8; Length 3264;
 Best Local Similarity 63.5%; Pred. No. 1.8e-191;
 Matches 2077; Conservative 0; Mismatches 1021; Indels 171; Gaps 8;

Qy 147 TGATCCGGGGCCACGAGACCCCAAGCCGCTGCGGGCGCTGAGCGGCGAGTGTGCGAGA 206
 Db |||||
 Qy 207 TATGCGGCGACGAGGTCCGCTACGGTGCAGCGCGACCTTCTCGTCCCTGCAACAGT 266
 Db |||||
 Qy 62 TCTGCGGCGACGGGTGGGACGACGCGGAGGGGCGAGCTGTCGCGCTCGACGCTCT 121
 Db |||||
 Qy 267 GCGGCTTCCCGGTGTCGCGCTCTGCTACGAGTACGAGCGCGGGAGGGGCGACGAGAACT 326
 Db |||||
 Qy 122 GCGGCTTCCCGGTGTCGCGCTCTGCTACGAGTACGAGCGCAAGGATGGCACCGGCTT 181
 Db |||||
 Qy 327 GCCCGGAGTGCAGACGCGCTACAGCGCTCAAGGGGAGCGGAGGTTGCCGGGAGCG 386
 Db |||||
 Qy 182 GCCCGGAGTGCAGACGCGCTACAGCGCTCAAGGGGAGCGGAGGTTGCCGGGAGCG 241
 Db |||||
 Qy 387 ATGACGAG-----GAGGACATCGACGACCTGGAGCAGAGTTCACATCG 431
 Db |||||
 Qy 242 AAGCGGAGTACTGATGCTGATGATGCTGATGATGATGATGATGATGATGATGATGATG 301
 Db |||||
 Qy 432 ACAGGAGTACTGAGAGGAGGAGTGGAGGGCAACATGCAAGACAGGAGTCAACGAGG 491
 Db |||||
 Qy 302 CCGACCAAGAGCAGAGATTGCTGATAGGATGCGAGTGGCGCATGAATGCTGGGGGTG 361
 Db |||||
 Qy 492 CGATGTGACGCGGAGGATGAGTACGGGAGGGGCGCGGAGCGGAGCGGCAACACA 551
 Db |||||
 Qy 362 GTGGAGACGTGCGGCGGTCACCAAGTATGACAGTGGCGAGATCGGGCTCACCAGTATGACA 421
 Db |||||
 Qy 552 CCGCGGAGATCCCGCCCATCATCAGCGCTCCCGTCCGTCGCGGTGAGCGGTGAGTTTC 611
 Db |||||
 Qy 422 GTGGGAGATCCCTCGGGGATACATCCCTTCACTAATAGCCAGATCTCGGGAGAA 481
 Db |||||
 Qy 612 CGATTACCAAGGGGTATGGCCACGGGAGGTCTGCTTCCCTGCAACAGCGCA----- 665
 Db |||||
 Qy 482 TCCCTGGAGCTTCCCTGATCATCATATGATGCTCTCTACCGGAAACATTTGGCAAGCGTG 541
 Db |||||
 Qy 666 -----TCCATCCGTACCTGTGCTGAGCCAGGAGTGCACAGTGGGACGAGAA 716
 Db |||||
 Qy 542 CTCATTCCCTTATGTAACCATTTACCAAAATCCATCAAGGGAGTCTCTGGGACGCAATTG 601
 Db |||||
 Qy 717 AAGAAGTGTAGCTGGAAGGAGGATGGACGACTGGAGTCAAGAGCGGCA----- 767
 Db |||||
 Qy 602 GAAATGTTGCTGGAAGAAAGAGTTGATGGCTGGAAACTGAAGCAGGACACAGGAGCA 661
 Db |||||
 Qy 768 -----TCTCGGCGCGC 779
 Db |||||
 Qy 662 TTCCCATGACCAACGGGACAAAGCATTTGCCCTTCTGAAAGTTCGGGAGTTGGTATATCG 721
 Db |||||
 Qy 780 GCGCGGATCCGGAAGACATGAGCGCGAGCTGGCATGTGAACGAGGAGCGGAGGACCGCG 839
 Db |||||

Db 722 ATGCATCCACTGATTACAAATATGAAGATGCCCTTACTGATGATGAACCTGCCAGCCTC 781
Qy
840 TGTGCGAGGAGGTGTCCGATCGATCGAGCAAGGTGAACCCGATCCGATGATCGTGG 859
Db
782 TCTCTAGAAAAGTTCCCTTCCTTCATCCAGATAAAATCCCTACAGAAATGGTCAATGTTTC 841
Qy
900 TGGCTCTGTTGTCGCGCTTCTTCTCCGTTACCGTATCCTTGCAACCCGCTCCCGGAGG 959
Db
842 TGGCTTGGTGTGTTAAAGCATTTTCTTCACTACCGTATTACGAATCCTGTGCGCAATG 901
Qy
960 CCATCGGCTGTGGCTGCTCCATCATCTGCAGATCTGGTTGCGCATCTCTGATCC 1019
Db
902 CGTATCGGCTCTGGCTTTTATCTGTTATATGTGAGATTTGGTTGTCTTGTCTGATAT 961
Qy
1020 TCGACCAAGTTCCCAAGTGGTTCCCATCGACCGCGAGAGTACTCGACCGCTCTCCC 1079
Db
962 TGGATCAGTTCCCGAAGTGGTTTCCAATCAACCGTGAACCTTACCTTGATAGCTGGCAT 1021
Qy
1080 TCAAGTACGAGAGGGAAGGGAGCGGTGGTGTCTGTGCGGCGTGAACCTGTTCTGAGCA 1139
Db
1022 TGAGGTATGACAGAGAAGGTGAGCCATCTCAGTTGGCTGTGTTGACATTTTGTGAGTA 1081
Qy
1140 CGGTGACCGCTCAAGGAGCGCGCTGTGTGACCGCAACACCGTCTCTCCATCCTCG 1199
Db
1082 CAGTCGACCCCATGAAGGAACCTTCTTCTGTACTGCCAATACCGTGTCTCAATCTTG 1141
Qy
1200 CCGTAGACTACCCGCTGGAAGAAGTCTCTGCTACGTCTCCGACGAGCGGCGTCCGATGC 1259
Db
1142 CTGTTGATTACCCAGTGGACAAGTCTCTTGTCTATGTATCTGACGATGGTGTGCAATGC 1201
Qy
1260 TGACGTTGAGTCGTCGTGCGGAGACGGCCGAGTTCCGGCGCAAGTGGTGGCCCTTCTGCA 1319
Db
1202 TGACTTTGATGATGGCTGAGACTTTCAGATTTGCTAGAAAAGTGGGTTCCATTCGTTA 1261
Qy
1320 AGAAGTTCCGATCGAGCCCGCGCCCGAGTTCTACTTCTCGCTCAAGGTCGACTACC 1379
Db
1262 AGAAGTATAACTTGAGCCGAAGAGCTCTCAGTGGTACTTCTCCGAGAAATCGATTACT 1321
Qy
1380 TCAAGGACAGGTGCGACCCACCTCTGTCAGAGAGCGCGCCATGAAGAGAGATATG 1439
Db
1322 TGAAGACAAAGTCCACCTTTCATTTGTTAAAGCCGTCTGTCATGAAGAGAGATATG 1381
Qy
1440 AGAGTTCAAGTCCGATCAACGCGTGTGGCCAAAGGCCATGAAGGTGCGCGGAGAGG 1499
Db
1382 AAGAAATCAAGATTAGGATAAATGGCTTGTGCTAAGGCACAGAAAAGTCCCGAGAGAG 1441
Qy
1500 GGTGGATCATGAAGGACGCAACGCTGGCCCGGGAACACACCCGCGACCAACCCCGGCA 1559
Db
1442 GATGGATCATGCAAGATGGCAACCATGGCCAGAAACAATATAGGGACCATCTTGBAA 1501
Qy
1560 TGATCCAGGTGTTCTGGGCCACAGCGGCGGCCACGACACCGAGGGCAACGAGCTGCCCC 1619
Db
1502 TGATTCAGGTTTCTTGGTCAAGTGTGGCTTGTATACCGAGGGTAATGAGCTTCCCC 1561
Qy
1620 GCCTGTGTACGTCTCCGTGAGAGAGCGCCCGGGAATTCAGACACCAAGAGGCGCGG 1679
Db
1562 GCCTGTGTACGTATCTGTGAGAAGCGTCTCGGTTTTCAGCACCAAGAGGCTGGTG 1621
Qy
1680 CCATGAACGCTGTGATTCGCTCTCCGCGTGTGACCAACGCGCATTCATCTCAACT 1739
Db
1622 CCATGAATGCTGTGTTGTTGTTCTCAGCTGTGCTTACCAATGGACAGTACATGTTGAATC 1681
Qy
1740 TGGACTGTGATCACTACATCAACAAACAGCAAGGCCATCCGGAGGGCCATGTCTTCTCA 1799
Db
1682 TTGATTTGATCACTACATCAACACAGCAAGGCTCTGCGGGAAGCTATGTGTTCTTCA 1741
Qy
1800 TGAACCTCAGTGTGCGCGGAAGGTCTGTACGTTTACGTTTCCGCGAGAGTTTCAACGCA 1859
Db
1742 TGGATCCAAACCTAGGAAGGAGTGTCTGTTATGTTTCAAGTTCCCAACAAGGTTCCGATG 1801
Qy
1860 TCGACGTGACGACCGATAGCTTAACAGGAACACCGTCTTCTTCGATCAACATGAAGG 1919
Db
1802 TTGATAGGAATGATCGATATGCGAAACAGGAACACTGTGTTTTTTCGATATTTAACTTTGAGG 1861

Qy 1920 GGCCTGACCGCATCCAAAGCCCGGTGTACGTGCGGACAGGTCGCTGTTCCGCGCGCAGG 1979
Db
1862 GTCTTGATGGCATCCAAAGGACCAAGTTTATGTGGAACTGTTGTGTATTTCAACAGAACTG 1921
Qy 1980 CCGTCTACCGCTACAAACCTCCCAAGGGACCCAGA----- 2015
Db 1922 CTCATATGTTTATGAACCCCAATTAAGCAGAGAAGAAAGGAAGTTCCTTGTCTATCAC 1981
Qy 2016 -----GGCCCAAGATGGTGACTGCTGCGCTGCTCCGCTGCTCCGCGCCCAAGAGCGGA 2069
Db 1982 TATGTGGGGCAGAAAGGCAAGCAAGTCAAAGAAAGAGAGTTCGGACAAGAGAGT 2041
Qy 2070 AACACGCCAA-----GGACGGGCTGCCCGAGGGCA 2099
Db 2042 CAAACAAGCAGCTGGACAGTGTGTGCCAGTTTTCATCTTGAAGATAAGAGAGGGTG 2101
Qy 2100 CCGCTGATATGGAGTAGATAGGCAAGAGATGCTCATGTCTCATGTCCCATCAATGAATTCGAGA 2159
Db 2102 TTGAAGGTCTGGATTCGATGATGAGAAATCACTTCTTATGTCTCAAATGAGCTTCGAGA 2161
Qy 2160 AGCGGTTCCGGCAGTCCGCGGCTTCGTCACTGACGCTGATGAGAGGAAGCGCGTCC 2219
Db 2162 AAGATTTGGCAGTCTGAGGTTTGTGCTCCTCACTCTCATGGAAATATGGTGGTTC 2221
Qy 2220 CTCCTTGTGCGACCCCGCGCTCTCAAAGAGGCCATCCATGTTCATCAGCTCGGGCT 2279
Db 2222 CTCATCTGCAACCCCAAGTCTCTTTGAAAGAGCTATCCATGTGATAAGTTGTGGCT 2281
Qy 2280 ACGAGCAAGACCCGACTGGGGCTGGAGC-----TGGGGTGAATCT 2321
Db 2282 ATGAGGACAAAGCCGAATGGGAACTGAGGCTATGCAACTTGCATGATGGGTGATCT 2341
Qy 2322 ACGGGTCGATCAAGAGGACATCTCTGACGGGTTCAAAGTTCGACTCGCGCGGTGGCGCT 2381
Db 2342 ACGGTTCCGTGACAGAGATATCTCCTGATTCAGATTCGATGCGCGAGGCTGAGAT 2401
Qy 2382 CCGTGTATCTGATGCGAAGCGGGCGGCTTCAAAGGGTGGCGCGCATCAATCTATCGG 2441
Db 2402 CAATCTACTGATGCTCAAGCGCCAGCTTCAAAGGGCTGTCTCTCATCAATCTTTTCAG 2461
Qy 2442 ACGGTTCAACCAAGTGTCTCGGTGGCGCTGGGTCCGCTCGAGATCTTCTTCAGCGGC 2501
Db 2462 ATCGCTTAAACCAAGTGTCTCGGTGGGCACTTGGTTCTGTGAAATTTCTTTTCAGTCGC 2521
Qy 2502 ACHAGCCCTGTGTACGCTACAGAAAGCGCAACCTCAAGTGGCTGAGAGCGCTTCGCT 2561
Db 2522 ATGTGCCATATGTTAGGCT---ATGGAGGACGCTTAACTTCTTGGAGAGATTTGCT 2578
Qy 2562 ACATCAACACCACTATTTATCCATTCGATCGATCCGCTTCTCATATATCTGTGTTTTC 2638
Db 2622 CCGCGTCTGCTCTCAACCGCAAGTTCATCATGCGCTCGATTCAGCATGCTTCGCTACTGCAACCTCC 2681
Qy 2639 CTGCTATCTGTTGCTCACTGGGAAGTTCATCATCCAGAGATTAGCAACTTTTCTAGTA 2698
Db 2682 TCTTCTTATCGCCCTCTTTCATGCTCATCTTTCGAGCGGCACTTCGAGAGATCGGTTGA 2741
Qy 2699 TTTGGTTATCTCTCTCTTCAATTTTGGCACTGGTATCCTTGGATGAGGTGGA 2758
Db 2742 GCGGGTGAAGCATCGAGGAGTGTGAGGAACGAGCAGTTCGTGGGTCACTGGCGGGTGT 2801
Qy 2759 GTGGTGTGGCATCGATGAGTGGGAGGAAAGAAAGTTCGGGTTATTTGGAGGTATAT 2818
Db 2802 CCGCGATCTTTCGCGGTGCTGAGGGCTGCTCAAGTTCCTCGCGGGATCGACACCA 2861
Qy 2819 CTGCGCATCTTTTGGCGTCTCCAGGGTCTCCTCAAGGTGCTTGTGTTATCGACACCA 2878
Db 2862 ACTTCAACCGTCACTCCAAAGGCCACCGGCGAGAGGACCAAGGATTCGCGAGGCTCTACG 2921
Qy 2879 ATTTCACTGTCACTCAAGGCTTCTGATGAGATGGCGAC---TTTGTGAGCTCTACA 2935
Db

QY 2922 CCTCAAGTGGACCAAGCTCTCTATCCCGCCACCAAGCTCTCTATCATTAACGTATCG 2981
 Db 2936 TGTTCAGTGGACCAAGCTCTCTATCCCGCCACCAAGCTCTCTATCATTAACGTATCG 2995
 QY 2982 GCGTCGTGGCGGATCTCCGACGCGCATCAACAAAGGCTACAGTCCCTGGGGGCCCTCT 3041
 Db 2996 GTGTTGTGTGGTATCTCATACGGATCAACAGGGCTACCATTCATGGGACGCTCT 3055
 QY 3042 TCGGCAAGCTCTTCTTGGCTTCTGGGTCATCGTCCACCTTACCGGTTCTCAAGGGGC 3101
 Db 3056 TTGGCAAGCTCTTCTTGGCTTCTGGGTCATCGTCCACCTTACCGGTTCTCAAGGGGC 3115
 QY 3102 TCATGGGGCGCCAGACGAGGACCCACCGTTGTGTGCATCTGTCCATTTCTGGCT 3161
 Db 3116 TTATGGGTGGCAAAACCGCACTCCGACCATCGTTGTGTGTGGCAATCTTCTGGCT 3175
 QY 3162 CCATCTTCTCTCTCTCTCTGGGTTCAGGATCGACCTTTTCATCGTCAGGACCAAGGCCCGG 3221
 Db 3176 CGATCTTCTCTCTCTCTCTGGGTTCAGGATCGACCTTTTCATCGTCAGGACCAAGGCCCGG 3235
 QY 3222 ACGTCAGGCAAGTGGGATCAATTCGCTGA 3250
 Db 3236 ATACCCAAACATGTGGCATCAACTGCTAA 3264

RESULT 14

AAZ99533
 ID AAZ99533 standard; DNA; 3704 BP.
 AC AAZ99533;
 XX
 XX
 DT 03-JUL-2000 (first entry)
 XX
 XX DNA encoding a maize cellulose synthase.
 DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW transgenic plant; plant breeding marker; ss.
 KW
 XX Zea mays.
 XX
 XX Key Location/Qualifiers
 FT 272..3496
 CDS /tag= a
 FT /product= "cellulose synthase"
 FT /transl_except= (pos: 1544..1546, aa: Xaa)
 FT /note= "no termination codon given; Xaa is an unspecified
 amino acid"
 XX
 XX W020009706-A2.
 PN
 XX
 XX 24-FEB-2000.
 XX
 XX 16-AUG-1999; 99WO-US018760.
 XX
 XX 17-AUG-1998; 98US-0096822P.
 PR
 XX (PION-) PIONEER HI-BRED INT INC.
 PA
 XX
 XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 PI
 XX
 XX WPI; 2000-224343/19.
 DR
 XX
 XX P-PSDB; AAY84121.
 DR
 XX
 XX New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.
 XX
 XX Claim 1; Page 191-196; 119pp; English.
 PS
 XX
 XX The present sequence encodes a maize cellulose synthase polypeptide. The
 CC cellulose synthase can be used for the improvement of stalk quality for
 CC improved stand or silage. It also provides an increased concentration of
 CC cellulose in the pericarp, hardening the kernel and improving its

CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting or
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein
 XX
 SQ Sequence 3704 BP; 909 A; 838 C; 962 G; 994 T; 0 U; 1 Other;

Query Match 35.1%; Score 1207.4; DB 3; Length 3704;
 Best Local Similarity 63.6%; Pred. No. 1.8e-190;
 Matches 2076; Conservative 0; Mismatches 1047; Indels 141; Gaps 10;

QY 155 GGCACAGAGACCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGTCGAGATATGCGGC 214
 Db 278 GCGCGGCGACCCACGAATTCGGGGAAGCATGTGCGCGGCGAGGTGCCAGATCGCGC 337
 QY 215 GAGGAGTGGGCTCACGGTGGAGCGGACTCTTCCTCGCTCAACGAGTGGGCTTC 274
 Db 338 GAGCGGTGGGCAACCGCGCGGAGCGGCACTCTTACCGCTCGACGTCTGCGGCTTC 397
 QY 275 CCGGTGTCGGCGCTCTACGAGTACGAGCGCGGGAGGGCAGCGAGAACTGCCCCAG 334
 Db 398 CCGGTGTCGGCGCTCTACGAGTACGAGCGCAAGGACGCGACCCAGGCGTCCCCGAG 457
 QY 335 TGCAGAGCGCTTACAGCGCTTCAAGGGAGCGCGGAGTTCGCGGGGAGCATGACGAG 394
 Db 458 TGCAGAACTAAGTACAGCGGCGCAAAAGGGAGGCCACACGTTACCGGTGAGGAAATGAG 517
 QY 395 GAGGACATCAGCAGCTGGAGCAGGTTCACATC-----GACGAGGAGAAATCAG 445
 Db 518 GATGTGATCTGACGATGTGAGTACTACACTACCAAGCATCTGGCAACCAAGATCAG 577
 QY 446 CAGAGGCGTGGAGGCAACATGCGAGA-----ACAGCCAGATCACCGAGCGATGCTGCA 501
 Db 578 AAGCAAAAGATTGTGAGAGAACTGCTCACTTGGCGGACAACTCACGTGGCAGTGATATT 637
 QY 502 CGGAGGAT-----GAGCTACGGGAGGGGCCCCGAGCGGCGAGCGGCAACACA 551
 Db 638 GCGCTGGCTAAGTATGACAGCGGTGAAATTTGGGCATGGGAAGTATGACAGTGTGAGATC 697
 QY 552 CCGCGCAGATCCCGCCCATCATCAGCGCT--CCCGCTCGCTGCGCGTGGAGTGT 609
 Db 698 CCTGTGGATATATCCCGTCACTAACTATAGCAGATCTCAGGAGAGATTCCTGGAGCT 757
 QY 610 TCGGATTACCAACGGGTATGCGCAACGCGGAGGTCTCGTCTTCTCCCTGCACAAAGCGCATCCA 669
 Db 758 TCCCTGATCATATGATGCTCTCTGTTGGGAACATTGGCAGGCGTGGACATCAATTTCT 817
 QY 670 TCCGT--ACCTGTGTCTGAGCCAGGAGTGCAGAGTGGGACGAGAAAGAGTGAAGC 727
 Db 818 TATGTTAAATCATTTCTCCAAACCCATCGAGGAGGTTCCTCCGTTAGCTTGGCAATGTTGCA 877
 QY 728 TGGAGGAGGAGGATGGACGACTGGAGTCCAA-----
 Db 878 TGGAAAGAGAGGGTGGATGATGGAATAAGGATAAAGGTGCAATTCCTATGACCAAT 937
 QY 760 -----GCAGGCGATCTCTCGCGCGCGCGCGCGATCCCGAA 793
 Db 938 GGAACAAGCATTCCTCCATCAGAGGCGGTGGAGTGTCTGATATTGATGCTTCTACTGAT 997
 QY 794 GACATGACCCCGACGTGGCACTGAACGACGAGGCGGAGCGCGCTGTTCGAGGAAGGTG 853

Db 998 TATAAATGGAATGCTTACTGAATGATGAACATCGGCAACCTCTATCTAGAAAGTG 1057
QY 854 TCATCGCGTCGAGCAAGGTGAACCCGTACCGGATGGTGTGATCGGTGGTGTGCTCTCGTTGTG 913
Db 1058 CCAATTCCTTCATCCAGATAAATCCGTACAGAAATGGTCATGTGTCTACGTTTGGCTGT 1117
QY 914 CTCGCCCTTCCTCCGCTACCGTATCCGTGCAACCCGTCCCGACGCCATCGCGGCTGTG 973
Db 1118 CTATGCATATTCCTTGGCTACCGTATCACACATCTCTGTGAACAATGCATATCCACTGTG 1177
QY 974 CTCGTCTCCATCATCTCGCGAGATCTGGTTCGCCATCTCTGGATCTCTGCACCGATTCGCC 1033
Db 1178 CTTTATCCGTCATATGTAGATCTGGTTCCTTGTCTGGATTTTGGATCATGATTCCTCA 1237
QY 1034 AAGTGTTCCTCCATCGACCGCGAGAGTGATCTCGACCGCTCTCCCTCAGGTCACGAGG 1093
Db 1238 AAGTGTTCCTCCATCAACCGTGAAATACATCTTGTAGACTGGCTTTAAGGTATGACCGA 1297
QY 1094 GAAGGGAGCCGTGCTGTCTCGCGGTGGACCTGTTCGTGAGCAGCGTGGACCCGCTC 1153
Db 1298 GAAGGTGAACCATCTCAATTAGCTCTCTGTGTATATTTTGTCAGTACTGTGGATCCAATG 1357
QY 1154 AAGGACCGCGTGTGACCGCAACACCGTCTCTCCATCTCTCGCGTAGACTACCC 1213
Db 1358 AAGGACCTCTCTGTACTGCAATACTGTCTTCCATCTCTGTCTGCGATTATCCG 1417
QY 1214 GTGGAAGGTCTCTGTACTGTCTCGACGCGCGCTCGATGCTGTGAGTTCGAGTGG 1273
Db 1418 GTTGACAAAGTATCTTGTATGTTCGGATGTGAGTGTCTATGCTGACTTTTGATGT 1477
QY 1274 CTGTGCGAGACGCCGAGTTCGCGCAAGTGGTGCCTCTTGCAGAAAGTTTCGGCATC 1333
Db 1478 CTCTCTGAACATTCAGAGTTTGTAGAAATGGTTCCGTTCTGTGAAGTACAACATA 1537
QY 1334 GAGCCCGCGCCCGGAGTCTACTTCTCGCTCAAGTGTGACTACCTCAAGNCAAGGTG 1393
Db 1538 GAGCCTTANGCCCGGAAATGGTATCTTGTCTCAGAAATTTGATTTACTTGAAGCAAAAGTT 1597
QY 1394 CAGCCACCTTCTGTGAGAGCGCCGCGCATGAAGAGAGATATGAGGAGTTTCAAGTTC 1453
Db 1598 CAACCTCATTTGTGAAGNACCGCGGCCATGAAGAGAGATATGAAGATTCAAAGTT 1657
QY 1454 CGATCAACCGCTGTGTGCAAGGCCAATGAAGTCCCGCAGAGGGGTGGATCATGAAG 1513
Db 1658 CGTATCAATGGTCTGTAGCAAGGCACAAAAGTTCCCGAGGAGGATGGATCATGCAA 1717
QY 1514 GACGCAACCGTGGCCCGGACACACCCGCGACACCCCGCATGATCCAGTGTTC 1573
Db 1718 GATGGTACACCTTGGCTTGGGAACAATACTAGGGACCATCTCGGAATGATTCAGGTTTC 1777
QY 1574 CTGGGCAACAGCGCGGCACAGCACCGAGGGCAACGAGCTGCGCCGCTCGTGTACGTC 1633
Db 1778 CTGGGTCAAGTGGAGGGCTTGACGTTGAAGGCATGAACCTCTCGTTGGTTATGTG 1837
QY 1634 TCCCGTGAGAGCGCCCGGATTCAGCACCAAGAAAGCCGCGCCATGAACGCTCTG 1693
Db 1838 TCTCGTGAATAAAGCTCTCTGATTTCCAACATCAAGAAAGGCTGGTCCCATGAATGCATT 1897
QY 1694 ATTCGGCTCTCCCGCTGTGACCAAGCGCCATTATGCTCAACTTGGACTGTGATCAC 1753
Db 1898 GTTCGTGTATCAGCTGTCTTACTAATGGGCAATACATGTGTGAATCTTGAATGTGACCAC 1957
QY 1754 TACATCAACAAAGCAAGGCCATTCGGGAGCCATGTGTCTCTCATGACCCCTCAGGTC 1813
Db 1958 TACATCAATATAGCAAGGCTCTTCGAGAAGCTATGTGCTTCTTATGAGACCCAAACCTA 2017
QY 1814 GGCCGAAAGGTCTGTCTACGTTTCAGTTCCCGCAGAGGTTTCGACGGATTCGACGTGACGAC 1873
Db 2018 GGAAGGAAATGCTGTGTATGTGCCAATTTCTTCAGAGGTTTGTATGATGAATGATGAC 2077
QY 1874 CGATCGCTTAACAGGAACACCGTCTTCTTCGATCAACATGAAGGGCTTGGACGGCATC 1933

Db 2078 CGATATGCAACAGGAACACTGTGTTTTTCGATATTAACCTTGAGAGGTCTTGACGGCAT 2137
QY 1934 CAAGGCCCGGTGTAGCTCGGACAGGGTGGTTCGCGCGCAGCGCTCTACGGCTAC 1993
Db 2138 CAAGGGCCAGTTTATGTGGAACTGGTGTGTGTAAACAGAACGGCTTATATGTTAT 2197
QY 1994 AACCTTCCCAAGGAGCCCAAGAGGCC-----CAAGATG 2026
Db 2198 GAGCCTCCAGTCAAGAAAAAAAGCCAGGCTTCTCTCTTCGCTTGTGGGGAAGAAA 2257
QY 2027 GTGACTGCGACTGTGCTCGCGTTCGCGCCGACAGAGCGGGAACACGCCAAGGAC--- 2083
Db 2258 AAGACGTCAAAATCTAAGAAGAGCTCGGAAAAAGAAAGTACATAGACGCGACAGAT 2317
QY 2084 -----GGGCTCGCGGAGGCGCACCGCTGATATGGAGTAGAT 2119
Db 2318 TCTGTACAGTATTTAACTCTCGAAGATATAGAGGAAGGATTTGAAGTTCTCAGTTTGT 2377
QY 2120 AGCGACAGGAGATGCTCATGTCCCAATGAATCTCGAAGAGGGTTCGGGCAAGTCCGG 2179
Db 2378 GATGAGAAATCGCTGATTTATGTCTCAAAATGAGCTTGGAGAAGAGATTTGGCCAGTCCAGT 2437
QY 2180 GCGTTCGTCACTCGACGCTGATGAGGAAGGGCGTCCCTCTTCGTGAGCCCGCC 2239
Db 2438 GTTTTGTAGCCTCTACTCTGATGAATATGGTGGTGTCCACAATCTGCAACTCCAGAG 2497
QY 2240 GCGCTCTCAAGAGGCCATCCATGTATCAGCTCGGCTAGGAGCAAGACCGACTGG 2299
Db 2498 TCTCTTCTGAAGAGACTATTCATGTATCATCAGCTGTGGCTATGAGGACAAAAGTGA 2557
QY 2300 GGGCTGGAGCTGGGTGGATCTACGGGTGATCAAGGAGGACATCTCTGACGGGTTCAAG 2359
Db 2558 GGAACCTGAGATTTGGGTGGATCTATGGTCTGTGTACAGAAGACATCTCACCGGATTCAG 2617
QY 2360 ATGCATGCGCGGGTGGGCTCCGTGTACTGATGCGAGCGGCGGGCTTCAAGGG 2419
Db 2618 ATGCATGCTGAGGGTGGCGATCAATCTGATGCGCTTAAGCGACAGCTTTCAAGGGA 2677
QY 2420 TCGCGGCCGATCAATCTATCGGACCGTCTCAACAGAGTCTCCGGTGGCGCTGGGGTCC 2479
Db 2678 TCTGCTCTATCAACTTTCGGATCGTTTGAATCAAGTGTCTTCGGTGGCTCTTGGTTCC 2737
QY 2480 GTGAGATCTTCTTACGCGGCAAGCCCCCTGTGTGATCGGCTACAAAGAACGGCAACCTC 2539
Db 2738 ATTGAAATTCCTTTTCAAGAGGCAATGTCCCATATGATATGGCT---ATGAGAGCCGCTT 2794
QY 2540 AAGTGGCTGGAGGCTTGGCTTACATCAACACCATCTACCCCTTCACTCGCTCCCG 2599
Db 2795 AAATTCCTGGAGAGATTTGCTTATATCAACAACAATTTATCCACTCAATCAATCCCG 2854
QY 2600 CTGCTCGCTACTGCAACCTCCCGCGTCTGCTCCCTCACCGGCAAGTTTCATATGCGG 2659
Db 2855 CTCCTCTGTACTGCATATTCGACAGAGTTTGTCTCTCAGTGGGAAGTTTCATATCCCA 2914
QY 2660 TCGATAGACGCTTCGCGAGCCTCTTTCATCGGCCCTCTTCATGTTCATCTTCGCGAGC 2719
Db 2915 AAGATTAGTAACTAGAGAGTGTGTGGTTTATATCGCTTCTTATCTCAATCTTTGCACT 2974
QY 2720 GGCATCTCGAGATCGGTGGAGCGGGTGAACATCGAGAGTGGTGGAGGAACGAGCAG 2779
Db 2975 GGTATCTTGAATGAGGTGGAGTGGTGTGGCAATGTATGAAATGGTGGAGGAACGAGCAG 3034
QY 2780 TTTCTGGGTCTATCGCGCGCTGTCGCGCATCTCTTCGCGCTCGTGAAGGGCTTCTCAAG 2839
Db 3035 TTTCTGGGTCTATTTGGTGTATTTCTGGCATTTATTTGGCGCTTTCAGAGGTCTCTCTGAAG 3094
QY 2840 GTCTCGCGCGGATCGAACCAATTCACCGTCACTCCAAAGGCCACCGCGGACGAGGAC 2899
Db 3095 GTGCTTGTGTGTATCGACACGAGCTTCACTGTCACTCTTAAGGCCAC---TGACGAAGAA 3151
QY 2900 GACGAGTTCGCGAGCTCTACCGCTTCAAGTGGACACCGCTCTCATCTCCGCGCCACCG 2959
Db 3152 GGTGATTTTGGCCAGCTCTACATGTTCAGTGGACACGCTTCTGTATCCACCACTACT 3211

QY 965 GGGCTGTGGCTCGTCTCCATCATCTCGAGATCTGGTTTCGGCATCTCCTGGATCTCGAC 1024
DB 1041 GCTTTTGGCTCATATCTGTATCTGTAATCTGGTTTGGCATCTTGGATTTCTTGAT 1100
QY 1025 CAGTTCCCAAGTGGTTCCCATCGACCGGGAGAGTACCTCGACGCCCTCTCCCTCAGG 1084
DB 1101 CAATTCCAAAGTGGTTCCCTATTGAGAGAGACTTACCTAGACCGGCTGTCACTAGG 1160
QY 1085 TAGCAGAGGAAGGGAGCCGTCTGCTGTGGCGGTGGACCTGTTTCGTGAGCAGCGTG 1144
DB 1161 TTGCAAGAAGGAGCCAGCCATCTCAACTTGCCTCAATTGATTTCTTTGTCACTAGCGTT 1220
QY 1145 GACCCGTCAAGAGAGCCGCGCTGGTGACCGGCCAAACACGCTGTCTCCATCTCGCCGTA 1204
DB 1221 GATCCCTTAAAGGAACCTCCTTTGGTCAACAATACTGTCTTATCTATCTCTTCGGTG 1280
QY 1205 GACTACCCCTGGACAGGTCTCTCTGTAGTCTCGACGACGCGGTGATGCTGACG 1264
DB 1281 GATTATCTGTGATAAGGTTCTTGTATGTTTCTGATGAGTGTCTCAATGCTTAACG 1340
QY 1265 TTGAGTCTGTTCGAGAGCGCGAGTTCGCGCGCAAGTGGGTGCCCTTCTGCAAGAAG 1324
DB 1341 TTTGAAGCATTAATCTGAAACATCTGAATTTGCAAGAATAAGGTTCTTCTGCAACGG 1400
QY 1325 TTGCGCATCGAGCCCGCGCCCGGAGTTCTACTTCTGCTCAAGTCTGACTACCTCAAG 1384
DB 1401 TACAATATTGAACCTCGCGCTCCAGAGTGTACTTCCAAACAGAAGATAGACTACTTGAAA 1460
QY 1385 GACAAGTTCAGCCCACTTCGTGACGAGCGCGCCGCTATGAAGAGAGTATGAGGAG 1444
DB 1461 GACNAGTGGCAGCAAACTTTGTAGGAGAGAGAGCAATGAAGAGAGTATGAGGAA 1520
QY 1445 TTCAAGTTCGGGATCAACGGGCTGGTGGCCAAAGGCCATGAAGGTGCCCGCAGAGGGGTGG 1504
DB 1521 TTCAAGGTGAGATCAATGCTTAGTTGCCAAGGCCAGAAAGTCTCTGAAGAAGATGG 1580
QY 1505 ATCATGAAGAGCGCAGCGTGGCCCGGGAAACAAACCCGCGACCAACCCCGCATGATC 1564
DB 1581 ACAATGAAGATGGAACCCCTGGCTGGAAACAAATTTGCTGATCATCTCGGAATGATT 1640
QY 1565 CAGGTGTTCTGGSCACAGCGGCGCCAGACACCGAGGSCAAGAGTGCCTCCGCTC 1624
DB 1641 CAGGTCTTCTTGGCCAAAGCGAGGCTTGACTGTGAGGGAATGAATGACCATG 1700
QY 1625 GTGTACGTCTCCGTGAGAAAGCGCCCGGGATTTCAGCACCAAGAGGCGCGCCATG 1684
DB 1701 GTTTATGTTCTAGAGAAACGACCAAGCTATAACCATATAAGAAAGCTGGTCTATG 1760
QY 1685 AACGCTCTGATTCGGCTCTCGCGTCTGACCAACCGCGCATTCATGCTCAACTTGGAC 1744
DB 1761 AATGCAATGGTCCGAGTCTCTGCTGTACTAACAAATGCTCCATATTTGTTAAACTTGGAT 1820
QY 1745 TGTGATCACTACATCAACAACGACCAAGCCATCCGGAGGCCATGTCTTCTCATGGAC 1804
DB 1821 TGTGATCACTACATCAACAACGACCAAGGCTATAAGGAAGCAATGTGTTTATGATGGAC 1880
QY 1805 CCTCAGGTCTGGCCGGAAGGTCTGTACGTTTCAAGTTCGCGCAGAGGTTTCGACGGCATCGAC 1864
DB 1881 CCTTTACTAGGAAGAAGGTTTGTATGTACAGTTTCCCTCAAGATTTGATGGATGAT 1940
QY 1865 GTGCAGACCGATACGCTAAACAGAAACACCGTCTTCTTTCGACATCAACATGAAGGGCTG 1924
DB 1941 CGCCATGACCGATATGCTAAACCGGAATGTTGTCTTTTTTGTATCAACATGAAGAAAGTTTG 2000
QY 1925 GACGGCATCAAGSCCGGTGTAGTCGGGACAGGTTGCTGTTCGGCGCCAGGCGCTC 1984
DB 2001 GATGGTATTAGGGTCCAAATTTATGTTGGTACTGGATGTGATTTAGAGCGAGGCATTA 2060
QY 1985 TAGCGCTACAAACCTTCCCAAGGAGCCCAAGAGGCC-----CAAGATG 2026
DB 2061 TATGTTTATGATGCCCCCAAAACAAAGAGCCACCATCAAGGACTTTCNACTGCTGCCCC 2120
QY 2027 GTGACCTGCGACTGCTGCTCGCGTCTTCGGCGGCAAGAAGCGGAAACACGCCAAGGACGGG 2086

DB 2121 AAGTGGTGTCTTTTGTGCTTGTGCTTTGGCAATAGGAAGCAAAAGAAAGACTACCAAAACC 2180
QY 2087 CTGCGCGAG----- 2095
DB 2181 AAAACAGAGAAAGAAAGATTATTTTTCAGAAAGAAAGAAACCATCCCTCGCATAT 2240
QY 2096 -----GGCACCGCTGATATGGGAGTAGAGACAGAGAGATGCTCATG 2140
DB 2241 GCTCTTGGTGAATTTGACGAAGCTGCTCCAGGAGCTGAGAAATGAAAGGCCGCTATTGTA 2300
QY 2141 TCCCAATGAACTTCAGAGAGCGGTTGGGCAAGTCCGCGCGTTCGTGCTGACGCTG 2200
DB 2301 AATCAACAATAAATTAGAAAAGAAATTTGGCCAAATCTTCTGTTTGTATACACACACTT 2360
QY 2201 ATGGAGGAAGGGCGCTCCTCTCTGTCGAGCCCGCGCTCTCAAGGAGGCCATC 2260
DB 2361 CTGGAATATGGTGGAAACCTTGAAGATGCAAGTCTCTCTTTTGAAGAAGCTATA 2420
QY 2261 CATGTCTCATGAGTTCGCGCTACGAGGACAAAGACACTGGGGGCTGGAGCTGGGGTGGATC 2320
DB 2421 CATGTCTATTAGTTGTGTTTATGAAGACAGACAGACTGGGAAAGAGATTGCTGGATC 2480
QY 2321 TAGGGTTCGATCAGGAGGACATCTTGACGGGTTCAAGATGCATCTGCGCGGGTGGCGC 2380
DB 2481 TATGGATCAGTTTACAGAGATATTCTAACTGGTTTCAAGATGATTTGTCTATGTTGGCGG 2540
QY 2381 TCCGCTGCTACTGCANTCGGAAGCGGGGCTTCAAGGGTTCGCGCGGATCAATCTATTCG 2440
DB 2541 TCAATTTACTGCTACACTAAACGGGTTGCATTCAAAGGTTCTGCACTCTGAAATCTTTCA 2600
QY 2441 GACGCTCTCAACAGAGTGTCTCGGTGGGGCTGGGGTTCGCTCGAGATCTTTCTTCAGCGG 2500
DB 2601 GATCGTCTTCCACAGGCTGCTCGGTGGGCTCTTTGGGCTTATTCAGATCTTCTTTCAGCAAT 2660
QY 2501 CACAGCCCCCTGCTGTACGGCTACAAGAACGGCAACCTCAAGTGGGTGGAGCGCTTCGCC 2560
DB 2661 CATTGGCCTCTTTGGTATGGGT---ATGGTGGCGGTCTGAAATTTTGGAAAGATTTTC 2717
QY 2561 TACATCAACACCACTACCTTACCTGCTCCGCTGCTGCTGCTTCTGCTTACTGCACCCCTC 2620
DB 2718 TACATCACTCCATCGTGTATCTTGGACATCTATTTCCTCTTGGCTTACTGTACATG 2777
QY 2621 CCGCGCTCTGCTCTCTCACCGCAAGTTTCATCATGCCGTCGATTAGCAGTTCGCCAGC 2680
DB 2778 CCTGCCATCTGTTTATTGACAGGAAATTTATCACTCCAGAGCTGAATAATTTGCCAGC 2837
QY 2681 CTCTCTTTCATCGCCCTCTTTCATGTCATCTTCGCAAGGGCATCTCTGGAGATGCGGTGG 2740
DB 2838 CTGTGGTTTCATGTCACCTTTTATCTGCAATTTTGTCTACGAGCATCTCTAGAAATGAGATGG 2897
QY 2741 AGCGGGTGAAGCATCGAGAGTGGTGGAGAACGAGCAGTTCTGGGTTCATTCGGCGCGCTG 2800
DB 2898 AGTGGTGTGGAATGATGATGCTGGAGAAATGACAGTTCTGGGTTCATTCGAGGTGTG 2957
QY 2801 TCOCGCTATCTCTTCCCGCTGCTGCAAGGGCTGCTCAAGGTCTCTCCCGGAGTTCGACACC 2860
DB 2958 TCCTCACACCTCTTCTGTTTCCAGGGACTTCTCAAGGTCTATAGCTGGTGTGATACA 3017
QY 2861 AACTTCACGCTCACTTCCAAAGGCCACCGGGAGCAGAGCAGAGTTTCGCGGAGCTCTAC 2920
DB 3018 AGCTTCACCGTGACATCAAAAG-----GGTGGAGATGATGAGGAGTTCTCAGAGCTATAT 3071
QY 2921 GCCTTCAAGTGAACACAGCTCTCTATCCCGCCCAACCAACGCTGCTCATATTAACGTCATC 2980
DB 3072 ACATTCAAATGAGTACTATTTGATACCTCTTACCACTTCTGCTTCTATTGAATCTATT 3131
QY 2981 GGCGTGTGGCGGCTCTCCGACGCCATCAACAACGGGTACCAAGTCTCTGGGGCCCCCTC 3040
DB 3132 GGTGTGGTGTGGCGTTTCAATGCGATCAATAACGGATATGAGTCAATGGGCGCCCCCTC 3191
QY 3041 TTCCGCAAGTCTCTTTCGCTTCTGGGTTCATGTCACCTCTACCGCTTCTCTCAAGGGG 3100

Db	3192	TTTGGGAAGCTATTCTTTTGCAATTTTGGGTGATTGTGCCATCTTTATCCCTTTCTCAAGGT	3251
Qy	3101	CTCATGGGGCGCCAGAAACAGGACGCCACCGTTGTTGTTCATCTGGTCCATTCTGTGGCC	3160
Db	3252	TTGGTTGGAAGGCAAAACAGGACACCAACGATTGTCTGCTGCTCCATTCTGTGGCT	3311
Qy	3161	TCCATCTTCTCCCTGCTCTGGGTGAGGATCGACCCCTTTTCATCTGTCAGGACCAAGGGCCG	3220
Db	3312	TCAATCTTCTGCTCCTTTGGGTTTCGGAATTGATCCTTTCTTGGGAAGGATGATGGTCCG	3371
Qy	3221	GACGTACGGCAGTGTGGCATCAATTGC	3247
Db	3372	CTTCTTGAGGAGTGTGTTGGATTGC	3398

Search completed: March 11, 2005, 19:16:34
Job time : 1207 secs

This Page Blank (uspto)

Db 548 GATCCGTGGGAGGAGGAGACGACACTGATGCGGATAGCGACTTCAATTAACCTTGCATC 607
Qy 433 CGACGGAATCAGCAGAGGAGCTGGAGGCAACATGCAGAAACAGCCAGATCA----- 485
Db 608 TGGCAATGAGGACCAGAGCAGAAGATTGCGGACAGAAATCGCGAGCTGGCGCATGAACGT 667
Qy 486 -----CCGAGGCGATGCTGCAAGGAGGATGAGCTACCGGAGGGGCCCCG----- 530
Db 668 TGGGGGACGCGGGGATGTTGGTCGCCCAAGATATGACAGTGGCGAGATCGGGCTTACCA 727
Qy 531 --ACGAGCGGACGCGCAACAAACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTC 588
Db 728 GTATGACAGTGGCGAGATTCCCTCGGGATACATCCCATCAGTCACTAACAGCCAGATCTC 787
Qy 589 CGTCCGGTGAGCGGTGAGTTTCC-GATTACCAACGGGTATGGCCACGGGAGGTCTCGT 647
Db 788 AGGAGAAATCCCTGGTGCTTCCCTGACCATATATGATGTCCCACTGGGAACATTGG 847
Qy 648 CTTCCCTGCAAGCGGATCCATCCGTAACCTGTGTGAGCAGGGAGTGCAGAGTGG 707
Db 848 CAAGCGTGTCTCCATTTCCCTATGTGAACCAATTCGCCAAATCCGTCAAGGGAGTTCTCTG 906
Qy 708 ACCGAGAAAGAGTGAAGTGGAGGAGGATGACGACTGGAAGTCCAAAGCAGGCA 767
Db 907 GTAGCATTTGGGAATGTTGCTTGAAAGAGAGGGTTGATGGCTGGAAATTAAGACGAGCA 966
Qy 768 -----TCC 770
Db 967 AGGGAGCATTCATGCAAGATGGCACAGCATTGTCCCTCTGAGGTTCGGGGTGTG 1026
Qy 771 TCGCGCGCGCGCGATCCGAGAGACATGAGCCGCGAGTGGCACTGAACGACGAGGCGA 830
Db 1027 GTGATATTGATGCATCAACTGATTACAAATGAGAAAGATGCTTTATTGAACGACGAAATC 1086
Qy 831 GGCAGCGCTGTCGAGGAGGTGTCGATCCGCTCGAGCAAGTGAACCCGTACCGGATGG 890
Db 1087 GACGCGCTCTATAGGAAGTTTCCACTTCTCTCCAGGATTAATCCCATACAGGATGG 1146
Qy 891 TGATCGTGTGCTCTCGTGTGCTCGCTTCTTCTCCGTTACCGGTATCCTGCACCCCG 950
Db 1147 TCATTGTGTCGATGATTGTTCTAAGCATCTTCTTGCACTACCGTATCACAATCTCTG 1206
Qy 951 TCCCGAGCGCATCGGGCTGTGCTGTCTTCATCATCTGCGAGATCTGTTGCGCATCT 1010
Db 1207 TGGGCAATGCATACCCATATTGCTTCTATCTGTATTATGTGAGATCTGGTTTGTCTTT 1266
Qy 1011 CCTGGATCCTCGACCACTCCCAAGTGGTTCCCATCGACCGCGAGAGTACCTCGACC 1070
Db 1267 CGTGGATATTGGATCAGTTCCCTAAGTGGTTTCCAATCAACCGGAGAGGTACCTTGATA 1326
Qy 1071 GCCTCTCCCTCAGGTACGAGAGGGAAGGGAGCGCTCGTGTGTCGGCGGTGGACCTGT 1130
Db 1327 GGCCTGGCATTAAGTATGACCGGAGGTGAGCCATCTCAGTTGGCTGTGTTGACATTT 1386
Qy 1131 TCGTGAAGCAGGTGGACCGCTCAAGAGCGCGCGTGTGACCGCCAAACACCGTCTCT 1190
Db 1387 TCGTCAGTACAGTCGACCCCAATTAAGAGGCTCTCTTGTCACTGCGCAATACCGTGTAT 1446
Qy 1191 CCATCTCGCGTGAAGTACCCCGTGGACAGAGTCTCTGCTACGTTCTCGAGCAGCGG 1250
Db 1447 CCATTTCTGTGGGAATACCCGTGGATTAAGGTCTCTTGTCTATGTATCTGATGATGGAG 1506
Qy 1251 CGTTCGATGCTGACGTTTCGAGTCCGTGCGAGACGCGCAGTTTCGCGCGCAAGTGGGTGC 1310
Db 1507 CTGCGATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTGTAGAAATGGGTAC 1566
Qy 1311 CTTTCTGCAAGAGTTTGGCATCGAGCCCGCGCCCGGAGTTTCTATTCTTCCTCAAGG 1370
Db 1567 CATTTGTTAAGAGTACAAATTAAGACTCTAGAGCTCTGAAATGGTACTTCTCCAGAAAA 1626
Qy 1371 TCGACTACCTCAAGGACAGGTGACGCCACCTTCGTGAGGAGCGCGCGCCCATGAGA 1430

Db 1627 TTGATTACTTGAAGGACAAAGTGCACCCCTTCAATTTGTTTAAAGACCGCGGGCCATGAAGA 1686
Qy 1431 GAGAGTATGAGGAGTTCAAGGTCGGATCAACCGCTGTGTGGCAAGGCCATGAAGGTGC 1490
Db 1687 GAGAAATGAAGAAATTCAAAGTTAGGGTAAATGSCCTTGTGTAAAGGCACAGAAAGTTC 1746
Qy 1491 CGGCAGAGGGTGGATCATGAAGGACGCAACCGCTGTGGCCCGGGAAACAACACCGCGGACC 1550
Db 1747 CTGAGGAGGATGGATCATGCAAGATGGCACCATGCGCCAGGAAACAATACCAGGGACC 1806
Qy 1551 ACCCGGCGATGATCCAGGTGTTCTCGGGCACAGCGCGCGCCACGACACCGAGGGGCAACG 1610
Db 1807 ATCTCGAATGATTCAGGTTTTCTTGGTCACAGTGTGGCTTGATCTAGAGGGCAATG 1866
Qy 1611 AGTCGCCCGCTCGTGTATACGTCTCCGTCGAGAGCGCCGGGATTCAGACACCAACAAGA 1670
Db 1867 AGCTACCCCGTTGGTCTATGTTCTCTGTAAGAGGTCTCTGGATTCAGCATCAACAAGA 1926
Qy 1671 AGCCCGGCGCATGAACGCTCTGAATTCGGCTCTCCGCGTGTGACCAACGCGCCCAATCA 1730
Db 1927 AAGCTGGTCCATGAATGCTCTTGTTCGTCTCTCAGCTGTCTTACCAATGGAACAATCA 1986
Qy 1731 TGTCTCAACTTGGACTGTGATCACTACATCAACAAACAGCAAGCCATCCGCGGAGCCATGT 1790
Db 1987 TGTGAAATCTTGATTTGATCACTACATTAACAACAGTAGGCTCTCAGGGAAGCTATGT 2046
Qy 1791 GCTTCTCTATGGAACCTCAGGTGCGCGGAAGGTCTGTCTACGTTCAAGTTCGCCGACAGGT 1850
Db 2047 GCTTCTTATGGAACCTTACCTAGGAAGGATGTCTGTCTACGTTCCAGTTTCCCAGAGAT 2106
Qy 1851 TGAACGCGATTCGACGTGCGACGATACGTAACAGGAACACCGTCTTCTTCGACATCA 1910
Db 2107 TCGATGGCATTCGACAGGAATGATCGATATGCCAAACAGGAACACCGTGTCTTCGATATTA 2166
Qy 1911 ACATGAAGGGTGGACGCAATCCAAAGGCGCGTGTGCTCGGACAGGAGTGGTGTTC 1970
Db 2167 ACTTGAAGGTCTTGATGGCATCCAAAGGACCAAGTTTATGTGCGAACTGGCTGTGTTTCA 2226
Qy 1971 GCGCCAGCGGTCTACGGCTACAACTCCCAAGGGAACCAAGA----- 2015
Db 2227 ACCGAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAGAGAGGGTGGTTTCTTGT 2286
Qy 2016 -----GGCCCAAGATGTTGACCTGCGACTGTGCGCGTGTTCGCGCGCAAGA 2063
Db 2287 CATCACTATGTGCGGTAGGAAGAGCAAGCAATCAAGAGAGGCTCGGACAAAGA 2346
Qy 2064 AGCGGAAACACGCCAAGGAC-----GGGCTGCCGAGG 2096
Db 2347 AGTCGAGAGAGCATGTGGACAGTTCTGTGCCAGTATTCACCTTGAAGATATAGAGGAGG 2406
Qy 2097 GCACCCCTGATATGGAGTAGATAGCAAGAGAGTGTCTCATGTCCCACATGAACTTCCG 2156
Db 2407 GAGTTGAAGCGCTGGATTTGACGACGAGAAATCACTTCTTATGTCTCAATGAGCTGG 2466
Qy 2157 AGAGCGGTTGGGCAAGTCCGCGGCTTGTCACTGCGAGCGTGTGATGAGGAAGGCGGCG 2216
Db 2467 AGAAGAGATTTGGCCAGTCCGAGCGTTTGTTCCTCCACTCTGTATGAGTATGTTGGTGTG 2526
Qy 2217 TCCCTCTTCTGTCGACCGCGCGCTCTCTCAAGAGAGCCATCCATGTCTCATGCTGCG 2276
Db 2527 TTCCCTCAGTCCGCACTCCGGAGTCTCTTCTGAAAGAAAGCTATCCATGTTTAAAGCTGTG 2586
Qy 2277 GCTACGAGACAAAGACCGACTCGGGCTGGAGCTGGGGTGGATCTCATCGGGTCCATCACCG 2336
Db 2587 GCTATGAGACAAAGACTGAATGGGAACTGAGATCGGGTGGATCTACGGTTCCTGTGACAG 2646
Qy 2337 AGACATCTGACGGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTGTACTGATGCTG 2396
Db 2647 AAGACATTTCTCAGCGGATTCAGAGTGCAGCGCGAGGCTGGCGGTCCGATCTACTGATGCTG 2706
Qy 2397 CGAAGCGGCGGGTTCAGGGGTTCGCGCGCATCAATCTATCGGACCGTCTCAACGAGG 2456
Db 2707 CCAAGCGGCGAGTTTCAAGGGGTCTGCCCCCATCAATCTTTTCGGAGCCGCTGTAACGAGG 2766

QY	891	TGATCGTGGTGGCTCTCGTGTGCTCGCTTCTTCTCCGGTACCGGTATCTCTGCACCCCG	950
DB	1147	TCATVTGCTCCGATGATGTTGTTCTAAGCATCTTCTTGCACTACCGTATCACAAATCCTG	1206
QY	951	TCCCGGACGCATCCGGCTGGCTCGTCTCCATCATCTCGAGATCTGGTTCGCCATCT	1010
DB	1207	TGCCAATGCATACCCATTATGGCTCTATCTGTATATGTAGATCTGGTTGCTCTTT	1266
QY	1011	CCTGGATCTCCAGCAGTTCCCAAGTGGTTCCCATCGACCGCGAGACGTACCTCGACC	1070
DB	1267	CGTGGATATTGGATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGGAGACGTACCTTGATA	1326
QY	1071	GCCTCTCCCTCAGTACAGAGGGAAGGGAGCCGTCGCTGCTGTCTGGCGGTGGAACCTGT	1130
DB	1327	GGCTGGCATTAAGGTATGACCGGGAAGTGAAGCCATCTCAAGTTGGCTGCTGTTGACATTT	1386
QY	1131	TCGTGAGCAGGTGACCCGCTCAAGAGCCCGCTGGTACCGCCCAACACCGTGTCT	1190
DB	1387	TCGTGATACAGTCAGCCCAATGAAGAGCCCTCTCTGTACATGCCCAATACCGTGTCTAT	1446
QY	1191	CCATCTCGCGGTAGACTACCCCGTGACAAGTCTCTCTAGCTCTCCGACGACGGCG	1250
DB	1447	CCATTCTGTGTGATTAACCTCTGTGATGAGTCTCTTGCTATGTATCTGATGTGAG	1506
QY	1251	CGTCTGATCTGACGTTCTGAGTCGCTGTCCGAGACGGCGGATTCGCGCGCAAGTGGGTGC	1310
DB	1507	CTGCGATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTGTGTAGAAAATGGGTAC	1566
QY	1311	CCTTCTGCNAGAAGTTCCGATCAGCCCGCGCCCGCGGAGTTCTACTTCTCGTCAAGG	1370
DB	1567	CAITTTGTAAGAAGTACAAACATGAACCTTAGAGTCTCTGATGGTACTTCTCCAGAAAA	1626
QY	1371	TCGACTACCTCAAGGACAAGTGCAGCCACCTTCGTGAGGAGCGCGCCGATGAAGA	1430
DB	1627	TTGATTTACTTGAAGGACAAGTGCACCTCTCATTTTGTAAAGCCCGCGGCCATGAAGA	1686
QY	1431	GAGAGTATGAGAGTTCAAGGTCGGATCAACCGCTGTGGTGGCAAGGCCCATGAAGTGC	1490
DB	1687	GAGAAATAGAAGATTCAAAGTTAGGTTAAATGGCCTTGTGTAAAGGCACAGAAAGTTC	1746
QY	1491	CGGACGAGGGTGTGATCATGAGACGGCAGCCGTCGCGCGGGGAAACACACCGCGACC	1550
DB	1747	CTGAGGAAGGATGATCATGCAAGATGGCACACCATGGCCAGGAAAACAATACCAGGGACC	1806
QY	1551	ACCCCGCATGATCAGAGTTCCTGGGCGACACGCGCGCACACACGAGGGCAAG	1610
DB	1807	ATCTCTGAAATGATTACAGTTTCTTGGTACAGTGTGGCTTGATCTGAGGGCAATG	1866
QY	1611	AGTGTCCCGCTCGTGTACGTCTCCGTGAGAGCGCCGGGATTCAGCACCAACAAGA	1670
DB	1867	AGCTACCCCGTTTGGTCTATGTTTCTCGTGAAAGCGTCTGGATTCAGCATCAACAAG	1926
QY	1671	AGCCGGGGCCATGAACGCTCTGATTCGCGTCTCCGCGTGTGACAAACGCGCATTC	1730
DB	1927	AAGCTGGTCCATGAATGCTCTTCTGGTGTCTCAGCTGTGCTTACCAATGGAATAACA	1986
QY	1731	TGCTCAACTTGGACTGTGATCACTACATCAACACAGAGGCCATCCGGAGGCCATGT	1790
DB	1987	TGTTGAATCTTGATTTGATCACTACATTAACAAGTAAAGGCTCTCAGGGAAGCTATGT	2046
QY	1791	GCTTCTCATGGAACCTCAGTCCGGCGGAAGTCTCTGCTACGTTTCAGTTTCCCGCAGAGT	1850
DB	2047	GCTTCTTATGGAACCTTAACTAGGAAGGAGTGTCTGCTACGTTCCAGTTTCCCGAGAT	2106
QY	1851	TCGACGGATCGACGTGACGACCGATACGCTAACAGGAACACCGTCTTCTTCGACATCA	1910
DB	2107	TCGATGGCAATTGACAGGAATGATCGATATGCCAACAGGAACACCGTGTTTTTCGATATTA	2166
QY	1911	ACATGAAGGGCTGGAAGGCATCCAAAGCCGGTGTACGTCCGGACAGGGTGGTGTTC	1970
DB	2167	ACTTGAGGGTCTTGTATGGCATCCAAAGACAGTTTATGTTCGGAACCTGGCTGTGTTTCA	2226

Db 1704 GAGAAATGAGAAATTCANAAGTTAGGGTAATAAGCCCTTGTGTGAAGGCAAGAAAGTTC 1763
Qy
Db 1491 CGGCAGAGGGTGGATCATGAAGAGCGCACGCCGTGGCCCGGGAACAACAACCGCGAAC 1550
Db 1764 CTGAGGAGAGATGATCATGAAGATGGACACCATGCGCCAGGAAACATACAGAGGACC 1823
Qy 1551 ACCCGGCATGATCAGAGTGTCTCTGGGCAACAGCGCGCGCCACGACACCGAGGGCAAG 1610
Db 1824 ATCTGGAATGATTCAGGTTTCTCTGGTCAACAGTGGTGGCTTGTGATCTAGAGGCAATG 1883
Qy 1611 AGTGGCCCGCTCGTGTAGTCTCCCGTGAGAGCGCCCGGATTCACACCAACAAGA 1670
Db 1884 AGCTACCCCGTTTGGTCTATGTTCTCTGTGAAGAGCGTCTGGATTCCAGCATCACAGA 1943
Qy 1671 AGGCGCGCATGAAAGCTCTGATTCGCGTCTCCGCGTGTGACCAACGCGCATTTCA 1730
Db 1944 AGCTGTGCGATGAAGTCTCTGTTCTGCTCAGCTGTGCTTACCAATGGACATPACA 2003
Qy 1731 TGCTCAACTTGGACTGTGATCACTACATCAACACAGCAAGGCGCATCCGGAGGCCATGT 1790
Db 2004 TGTGAACTTGAATTGTGATCACTACATTAACACAGTAAAGCTCTCAGGGAAGCTATGT 2063
Qy 1791 GCTTCCTCATGAGCCCTCAGTCCGCGGAGAGTCTGCTAGTTTCAAGTCCCGCAGAGGT 1850
Db 2064 GCTTCCTTATGGACCCCTAACTAGGAAGAGTGTCTGCTACGTCAGTTCCTCCCGCAGAT 2123
Qy 1851 TCGAGGCGATCGACGTCGACGCGATACGCTTAACAGGAACACCGTCTCTTCTGACATCA 1910
Db 2124 TCGATGCGATGACAGGAATGATCGATATGCCAACAGGAACACCGTGTGTTTTCATATTA 2183
Qy 1911 ACATGAAGGGCTGGACGGCATCAAGCGCGGTGTACGTCGGGACAGGGTGGTGTTC 1970
Db 2184 ACTTGAGAGTCTTGATGGCATCAAGGACAGTTCATGTCCGAACTGGGTGTTTCA 2243
Qy 1971 GGGCCAGGGCTCTAGCGCTACAACTCCCAAGGAGCCCAAGA----- 2015
Db 2244 ACCGAACAGCTCTATATGTTGTATGAGCCCCCAATTAAGCAGAAAGGGTGGTTCCTGT 2303
Qy 2016 -----GGCCCAAGATGTGACCTGCGACTGTCGCCGTGCTCGCGCGCAAGA 2063
Db 2304 CATCACTATGTGGCGGTAGGAGGAAGCAAGCAATCAAGAGGGCTCGGACAGAGA 2363
Qy 2064 AGCGAAACAGCCCAAGAC-----GGGCTGCCGAGG 2096
Db 2364 AGTCGAGAAGCATGTGACAGTTCTGTGCGAGTATTCAACCTTTGAAGATATAGAGGAG 2423
Qy 2097 GCACCGCTGATATGGGAGTAGACGACAGGAGATGCTCATGTCCCATGNACTTCG 2156
Db 2424 GAGTTGAAGCGCTGGATTTGACGACGAGAAATCACCTTCTATGTCTCAAAATGAGCCTGG 2483
Qy 2157 AGAAGCGGTTCCGGCAGTCCGCGCGTTCGTCACTCGAGCGCTGATGGAGGAAGGCGCG 2216
Db 2484 AGAAGATTTGGCCAGTCCGCGCGTGTGTGCTCCACTCTGATGGAGTATGGTGGTG 2543
Qy 2217 TCCCTCTCTGTGAGCGCGCGCGCTCTCAAGAGGAGGCATCCATGTATCAGTGTGCG 2276
Db 2544 TTCCTCAGTCGCAACTCCGAGTCTCTTCTGAAAGAAAGCTATCCATGTTTATAAGCTGTG 2603
Qy 2277 GCTACGAGGACAGACGACTGGGGGTGGAGCTGGGGTGGATCTACGGGTCCGATCAGG 2336
Db 2604 GCTATGAGGACAAGACTGAATGGGGAATGAGATCGGGTGGATCTACGGTTCGTGTGACG 2663
Qy 2337 AGGACATCTGACCGGGTTCAAGATGCACCTGCCCGGTGGCGTCCGTGTACTGTGATGC 2396
Db 2664 AAGACATCTCACCGGATTCAGATGACGCGCGAGGCTGGCGTCCGATCTACTGCTATGC 2723
Qy 2397 CGAAGCGGGCGGCTTCAAGGGGTCCGCGCGCGATCAATCTATCGGACCGTCTCAACGAG 2456
Db 2724 CCAAGCGGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTCCGAGCGGTCTGAACGAG 2783
Qy 2457 TGCTCCGGTGGCGCTGGGTCCGTCGAGATCTTCTCAGCGGACAGCCCGCTGCTGT 2516

Db 2784 TGCTCCGGTGGGCTCTTGGGTCCGTGAGATCTCTTTACGCGGCACTGCCCCCTGTGGT 2843
Qy 2517 ACGGCTTACAAGAACGGCAACCTCAAGTGGCTGAGCGCTTCGCTATCATCAACACCAACCA 2576
Db 2844 ACGGCTTAC---GGAGGGCGGCTCAAGTTCCTGGAGAGATTCGGGTATCATCAACACCA 2900
Qy 2577 TCTAACCCCTTCACTCCGCTCCGCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2636
Db 2901 TCTAACCCGCTCAGTCCATCCGCTTCTCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTG 2960
Qy 2637 TCACCGGCAAGTTCATCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2696
Db 2961 TCACCGGAAAGTTCATCATCTCCAGAGATCAGCAACTTCGCGCAGATCTGTTTCACTTCCC 3020
Qy 2697 TCTTCACTGCTCACTTTCGCGACCGGCTCCTGAGAGATGCGGTGGAGCGGGTGGAGCATCG 2756
Db 3021 TCTTCACTGCTGCTTTCGCGACCGGCTCCTGAGAGATGAGTGGAGGGGTGGGATCG 3080
Qy 2757 AGAGTGGTGGAGGAACGAGCAGTTCCTGGGTCTATGCGGCGCGTGTGCGGCTGCTCTTTCG 2816
Db 3081 ACAGTGGTGGAGGAACGAGCAGTTCCTGGGTCTATGCGGGGCGCATCTCCGCGCAGCTTTCG 3140
Qy 2817 CGGTGCTGAGGCGCTGCTCAAGTCTCTCCGCGGATCGACACCACTTCACCGTCACTT 2876
Db 3141 CGGTGTTCCAGGGCTGCTCAAGTGTGCGCGGCTCGACCACTTCACCGTCACTT 3200
Qy 2877 CCAAGCGCACCGGCGACGAGGACGAGTTCGCGGAGCTCTACGCGCTTCAAGTGGACCA 2936
Db 3201 CCAAGGCTCG---GACGAGGACGGCGACTTCGCGGAGCTGTATGTTTCAAGTGGACGA 3257
Qy 2937 CGTCTCTCATCCCGCGCACGCTGCTGCTCATGTTAAAGTCACTGCGGCTGCTGCGCGGCA 2996
Db 3258 CGTCTCTGATCCCGCGCACCCACCATCTGATCATCAACTGCTGCGGCTGCTGCGCGGCA 3317
Qy 2997 TCTCCGAGCGCATCAACAGCGGTACAGTCCCTGGGGGCGCTTCTGGCAAGCTCTTCT 3056
Db 3318 TCTCTACGGCATCAACAGCGGATACAGTCTGGGGCGCGCTTCTGGCAAGCTCTTCT 3377
Qy 3057 TCGCCTTCTGGGTCTACGTCACCTCTACCGGCTTCTCAAGGGGCTCATGGGGCGCGAGA 3116
Db 3378 TCGCCTTCTGGGTCTACGTCACCTCTACCGGCTTCTCAAGGGGCTCATGGGGCGCGAGA 3437
Qy 3117 ACAGGAGCGCCAGCTGTTGTTGCTCATCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTG 3176
Db 3438 ACCGCAACCGGACATCGTCTGCTGCGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 3497
Qy 3177 TCTGGGTGAGATCGACCTTTCATGTCAGGACCAAGGGCGCGGAGCTCAGGCGAGTGTG 3236
Db 3498 TGTGGGTGCGATCGACCCCTTCAACACCGCGCTCACTGGCGCGGATACCAAGAGTGTG 3557
Qy 3237 GCATCAATTGCT 3248
Db 3558 GCATCAACTGCT 3569

RESULT 4

US-10-160-719A-25
; Sequence 25, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Heilentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,363
; PRIOR FILING DATE: 1999-08-06

Qy	1960	GTGCGTGTTCGGCGCCAGGCGCTCTACGGCTACAAACCTCCCAAGGGAACCCCAAGAGGCC	2019
Db	2131	ATGTGTTTTTCAGACGGCAGGCACTGTATGGTTATGATGCTCTCTAAACGAGAAGCCACCC	2190
Qy	2020	-----CAAGATGGTGACCTGGGACTGCTGCCGTGCTTCCGCGCGCAAGAGCG	2067
Db	2191	ATCAAGAACTTGGCAACTGCTGGCCCAAGTGGTGGCTCTCTTGTGCTGCGACGAGAACAA	2250
Qy	2068	GAACACGCCCAAGGACGGGTGCCGGAG-----	2095
Db	2251	GAATAAAGAGACTACAAACCAAAGACGGAGAGAGAAAGATATATTTTTCAGAA	2310
Qy	2096	-----GGCACCGCTGATATGGGAGT	2115
Db	2311	AGCAGAAACCCATCTCTGCATATGCTTTGGGTGAAATTGATGAAGGTGCTCCAGGTGC	2370
Qy	2116	AGATAGCGACNAGAGATGCTCATGTCCCACTGAACCTTCGAGNAGCGGTTCGGGCAATC	2175
Db	2371	TGATATCGAAGAGCGCGGAATCGTAAATCAACAGAAACTAGAGAAGAGAAATTTGGGCAATC	2430
Qy	2176	CGCGGCGTTTGGTCACTGCTGACGCTGATGGAGGAAGCGGCTGCCCTCTTCTGCTCGAGCCC	2235
Db	2431	TTCTGTTTTTGTGCGCATCAACATCTCTTGAGAACGGAGGAGCCCTGAAGAGCGCAAGTCC	2490
Qy	2236	CGCCGCGCTCCTCAAGGAGGCCATCCATGTTCATCAGCTGGCGGTACGAGGACCAAGACCGA	2295
Db	2491	AGCTTCTCTTCTGAAGGAAGCTATACATGTTATCAGCTGCGGTCTACGAACAGACCGA	2550
Qy	2296	CTGGGGGCTGGAGCTGGGGTGATCTACGGGTGATCACGAGAGACATCTCTGACGGGTT	2355
Db	2551	CTGGGGAAGAGAGATTGGCTGGATTTACGGATCGATCACAGAGGATATCTTGACTGGATT	2610
Qy	2356	CAAGATCACTGCGCGGCTGGCGCTCCGCTGCTACGATCCGAGCGGGGGGGGTTCAA	2415
Db	2611	TAAGATGCATCGCCATCGCTGGCGGTCTATTTACTGCTATCCGAGCGGCTCGATTCAA	2670
Qy	2416	GGGTCGCGCGCCGATCAATCTATCGGACCGTCTCAACAGAGTGCTCCGGTGGGCGCTGGG	2475
Db	2671	AGGTTCTCGCGCTCTGAACCTTTCGACCGCTCTTCACAGGTCTTCGCTGGGCGCTTGG	2730
Qy	2476	GTCGCTCGAGATCTTCTTCAGCGGCGACAGCCCGCTGCTGACGGCTACAAAGAACGGCAA	2535
Db	2731	GTCGCTCGAAATTTCTTCAGCAAGCACTGCCCACTTTGGTACGGATAC--GGCGGCGG	2787
Qy	2536	CCTCAAGTGGCTGGAGCGCTTCGCTATACATCAACACCAACCATCTACCCCTTCACCTCGCT	2595
Db	2788	GCTAAATTTCTGAAAAGGTTTCTTATATCAATCCATCCCTGTTATCCCTTGGACGTCCAT	2847
Qy	2596	CCCGCTCTCGCTTACATGCAACCTCCCGCGCTCTGCTCTCAACGGCAAGTTTCATCAT	2655
Db	2848	TCCTCTCTCGCTTACTGTACCTTGGCTGCGCATCTGCTGCTCAAGGGGAAGTTTATCAC	2907
Qy	2656	GCGCTGATATAGCATGTTTCGCCAGCTCTTCTTCATCGCCCTCTTCATGTCATCTTCGC	2715
Db	2908	ACCAGAGCTTACCAATGTGCGCAGTATCTGGTTTCATGGCACTTTTCATCTGCATCTCCGT	2967
Qy	2716	GACGGGCAATCTTGAGATGCGGTGGAGCGGGGTGAGCATCGAGAGTGTGTGAGGAAACGA	2775
Db	2968	GACCGGCATCTTGAATATGAGGTGGAGTGGCGGTGCCATCGACACTGTGTGAGGAAACGA	3027
Qy	2776	GCAGTTCTGGGTCAATCGGCGCGTGTCCGCGCATCTCTTCGCGTGTGTGAGGGCGCTGCT	2835
Db	3028	GCAGTTCTGGGTCAATCGAGGCGTTTTCGGCGCATCTGTTTCGCGTGTTCAGGGGCTTGCT	3087
Qy	2836	CAAGGTCTCTCGCGGATCGACACCAACTTTCACCGTCACTCCCAAGGCCACCGCGCACGA	2895
Db	3088	GAAGGTGTTTCGCGGCATCGACACGAGCTTCAACGTGACGCTCGAAGGCCCGGGGACGACGA	3147
Qy	2896	GGACGACAGTTTCGCGGAGCTTACGGCTTCAAGTGGACCAACGCTCTCTCATCCCGCCAC	2955
Db	3148	G-----GAGTTCCTCGAGCTGTACAGTTCAAGTGGACCAACCTGCTGTGATCCCCCGAC	3201

RESULT 5

US-10-160-719A-45

US-10-100-715A-43
: Sequence 45, Application US/10160719A

Patent No. 6803498

: GENERAL INFORMATION:

APPLICANT: Dhugga, Kanwarpal S.

APPLICANT: Helentjaris, Timothy G.

APPLICANT: Bowen, Benjamin A.

APPLICANT: Wang, Xu

TITLE OF INVENTION: Maize Cellulose Synthases and Uses

; TITLE OF INVENTION: Thereof

FILE REFERENCE: 0864C

; CURRENT APPLICATION NUMBER: US/10/160,7

;
CURRENT FILING DATE: 2002-06

; PRIOR APPLICATION NUMBER: US 60/096,822

; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: US 09/371,383

;
PRIOR FILING D

; NUMBER OF SEQ

SOFTWARE: Fast

```

; SEQ ID NO 45
      LENGTH: 2013

```

```

; LENGTH: 3813
: TYPE: DNA

```

TYPE: DNA
ORGANISM: Zea mays

```

; ORGANISM: zea mays
: FEATURE:

```

FEATURE: NAME/KEY: CDS

Query Match 39.6%; Score 1364.8; DB 4; Length 3813;
Best Local Similarity 65.7%; Pred. No. 1.7e-243;
Matches 2202; Conservative 0; Mismatches 1007; Indels 143; Gaps 9;

Qy	30	CTGTCGAGCACCTGAGGGGTCGGAGGCGCGAGCTAGCCTAGCACGCGGGCCTTCGCGCG-	88
Db	151	CTGTGCGCGCCGCGCGGGGTCGTGCGCAGCGAGATCCGCGGGGCGGGGCGGGGGGCC	210
Qy	89	-GCATGGAGGCCACGAGCCGGGCTGTGTGGCCGGCTCGCACAAACGGGAACGAGCTGGTGCT	147
Db	211	TGAGTGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATAAACGGGAACGAGCTGGTGCT	270
Qy	148	GATCCGCGGCGCA-CGAGGA CCCCAGCGCTGCGGGCG------CTGAGCGGGCA	195
Db	271	GATCCGCGCGACCGCGAGTCGGAGCGCGGGCGGGCGCGCGCGCGCGCGCGGCGAGC	330
Qy	196	GGTGTGCGAGATAGCGCGACAGAGTCGGCTCACGTGACACGCGACCTCTTCGTCGC	255
Db	331	GCCGTGCAGATATGCGGCGACAGGTCGGGGTGGGCTTCACCGGGGAGCCCTTCGTGCG	390
Qy	256	CTGCAACGAGTGGGCTTCCCGGTGTCGGGCCCTGCTACGAGTACGAGCGCCGGAGGG	315
Db	391	GTGCAACGAGTGCSCCTTCCCGTCTCCGGCGCTGTACGAGTACGAGCGCGCGAGGG	450

Db 2611 TAGATGCACTGCATGGCTGGCGTCTATTACTGCAATCCCAAGCGGCTGCATCAA 2670
QY 2416 GGGGTGGCGCCGATCAATATATCGAACCGTCTCAACAGGTGTCGCGTGGCGCTGGG 2475
Db 2671 AGGTTCCTGGCCCTCTGAACCTTTCGACCGTCTTCCACAGTCTTTCGCTGGCGCTTG 2730
QY 2476 GTCCGTGCAATCTTCTTACGCGGACAGCCCGCTGCTGTACGGCTACAGAAACGGCAA 2535
Db 2731 GTCCGTGCAAAATTTCTTACAGCAAGCACTGCCACTTTGTACGGATAC---GGCGGCG 2787
QY 2536 CCTCAAGTGGCTGAGCGCTTCGCTACATCAACACCAACCATCTACCCCTTCACTCGCT 2595
Db 2788 GCTAAATTCCTGAAAGGTTTCTTATATCAATCCATCGTTATCCCTGGACGTCAT 2847
QY 2596 CCGCTGCTCGCTACTGCACCCCTCCCGCGCTGCTGCTCTCACCGGCAAGTTTCATCAT 2655
Db 2848 TCCTCTCTGGCTTACTGTACTTGCCTGCCATCTGCTGCTCACGGGGAAGTTTATCAC 2907
QY 2656 GCGGTGATAGACAGTTCCGCCAGCTCTTCTTCAATGCGCTCTTCAATGTCATCTTCGC 2715
Db 2908 ACCAGAGCTTACCAATGTCGCAAGTATCTGGTTCATGGCACTTTTTCATCTGCATCTCGT 2967
QY 2716 GACGGGCATCTGGAGATGCGTGGAGCGGGGTGAGCATCGAGAGTGGTGGAGAACGA 2775
Db 2968 GACCGGCATCTGGAATAGGTGGAGTGGCGTGCGCCATCGACACTGGTGGAGAACGA 3027
QY 2776 GCAGTTCCTGGCTCATCGCGCGCTGTCGCGCATCTCTTCGCGCTGCTGTCAGAGGCTGCT 2835
Db 3028 GCAGTTCCTGGCTCATCGAGGCGTTCGCGCATCTGTTTCGCGTGTTCAGAGGCTGCT 3087
QY 2836 CAAGTTCCTCGCGGATCGACACCACTTCAACCGTCACTTCAAGGCCACCGCGACGA 2895
Db 3088 GAAAGTGTTCGCGGCATCGACACGAGCTTCAACCGTCACTTCAAGGCCCGCGACGA 3147
QY 2896 GACGACGAGTTCGCGAGCTTACGCTTCAAGTGGACACCGCTCTCATCCGCGCCAC 2955
Db 3148 G-----GAGTTCGAGCTGTACACGTTCAAGTGGACACCGCTCTGTATACCCCGAC 3201
QY 2956 CACGCTGCTCATATTAACTGATCGCGCTGCTGGCGCGATCTCCGACCGCCATCAACAA 3015
Db 3202 CACGCTCTCTGCTGNACTTCATCGGGTGTGGCGCGGATCTGNAACGCGATCAACAA 3261
QY 3016 CGGGTACCACTCTCGGGGCGCTTTCGGAAGCTTCTTTCGCTTTCGGGTCAATGCT 3075
Db 3262 CGGGTACGAGTCTGGGGCGCTTTCGGAAGCTTCTTTCGCTTTCGGGTGATGCT 3321
QY 3076 CCACCTTACCGTTCCTCAAGGGCTCATGGGGCGCCAGAACGAGCGCCACCGTGT 3135
Db 3322 CCACCTGTACCGGTTCCTCAAGGGTCTGGTGGGAGGACAGAACGAGCGCCGATGCT 3381
QY 3136 TGTCACTGCTGCTCATCTGCTGGCTCCATCTTCTCCTGCTCTGGGTCAAGATCGACCC 3195
Db 3382 CATCGTCTGCTCATCTGCTGGCTCGATCTTCTGCTCTTGGGTGCGGTGCAACC 3441
QY 3196 TTTCACTGTCAGACCAAGGCGCGGACGCTCAGCGAGTGTGGCATCAATTCG 3247
Db 3442 GTTCTCTCGCAAGACAAAGCGCGCTCTCGGAGGAGTGTGGCTGGATGCTG 3493

RESULT 6

US-10-160-719A-1
; Sequence 1, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A

; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)...(3239)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3487)...(3487)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(3568)
; OTHER INFORMATION: n = A,T,C or G
US-10-160-719A-1

Query Match 35.5%; Score 1222.4; DB 4; Length 3568;
Best Local Similarity 63.5%; Pred. No. 3.9e-217;
Matches 2033; Conservative 0; Mismatches 1051; Indels 120; Gaps 6;
QY 155 GGCACGAGACCCCAAGCGCTGCGGGCGCTGAGCGGCGAGGTGTGCGAGATATGCGGC 214
Db 42 GCGCGCAGGGCCGAGCCCATGACAGCGGAACGGCGAGGTGTGCCAGATTGCGGC 101
QY 215 GACAGGTTCGGGTCTACGTGGACGGCGACTTTCGTTCGCTCGCTGCAACGAGTTCGGCTTC 274
Db 102 GACGACGTGGGGCGCAACCCCGACGGGAGCCCTTCGTGGCTGCAACGAGTTCGGCTTC 161
QY 275 CCGGTGTGCGGCGCTGCTACGAGTACGAGCGCGGGAGGGGACGAGAACTGCCCCCAG 334
Db 162 CCATCTGCGGAGCTGTCTACGAGTACGAGCGCGCGGAGGCGACGAGAACTGCCCGCAG 221
QY 335 TGCAGACCGCTTACAGCGCTCAAGGGGAGCCCGAGGTTGCGCGGGACGATGACGAG 394
Db 222 TGCAGACCGCTTCAAGGGTGTGCGCGCGCTGCGCGGGGACGAGGAGAG 281
QY 395 GAGGACATCGACGACTGAGCAACGAGTTCAACATCGACGAGAGAAATCAGCAGAGCAG 454
Db 282 GACGGCTGCAACGACTGAGAAACGAGTTCAACTGGAGGCAACGAGCAGCACTCCAGTAC 341
QY 455 CTGGAGGCAACATGACAGAACGACAGT-----CACCAGGCG 493
Db 342 CTCGCGAGTTCATGTCTCACGCGCCCATGAGCTACGGCGCGGCGCGACCTCGACGCG 401
QY 494 ATGCTCAGCGGAGGATGAGCTACGGGAGGGGCGCGAGCGAGCGGCGCAACACACC 553
Db 402 GTGCGCAGCCATTCACCCCATGCCAATGTTCCCTCTCTACCAACGAGCAGATGGTC 461
QY 554 CCGCAGATCCCGCCCATCATCACCGGCTCCCGCTCCGCTCCGCTGAGCGGTGAGTTTCG 613
Db 462 GATGACATCCCGCGGACGAGCAGCGCTTGTGCGCTCTGTTGCGGTGCGCGGGGAG 521
QY 614 ATTACCAACGGGTATGGCCACGGCGAGGTCTGCTTCTCCTGACAGGCGCATCCATCCG 673
Db 522 AGGATTCACCTCTCCCGTACCGGATCCCAACCTTCTGTCACACCGAGGTCTATGGAC 581
QY 674 TACCTGTGTCTGAGCCAGGAGTGCACAGTGGGAGGAGAGAAAGAAAGTGAAGTGAAG 733
Db 582 CCTTCAAGATCTCCCGCATATGCTACGGGAGGTAGCATGGAAGAGAGATGGAG 641
QY 734 GAGAGATGAGACACTGGAAGTCCAAAGCAGGGCATCTTCGGCGGCGCGCCGATCCGAA 793
Db 642 AGCTGGAAGCAGAAAGCAGGAGGAGATGCACACAGAGGAAACGATGGCGCGCGATGAT 701
QY 794 GACATGAGCGCGACGTGGCACTGAACGAGGCGGAGGCGAGCGCTGTTCGAGGAGGTG 853

Db 702 GGTGATGATGAGATCTACCACTAATGGATGAAGCTAGACAGCCATTTGTCAGAAAGATC 761
Qy 854 TCGATCGCGGTGAGCAAGGTGAACCCGTACCGGATGGTGTGATGCTGGTCTCGTGTG 913
Db 762 CCGCTTCTTCAAGCCAAATCAACCCCTATAGAGTGAATTATAAATAATTCGGCTAGTGGT 821
Qy 914 CTCGCTTCTTCTCCGCTACCGTATCTGCAACCCGTCGCGAGCCATCGGCTGG 973
Db 822 TTGTGTTTCTTCCACTACCGAGTGATGATCCGGTGCCTGATGCAATTTGCTTTATGG 881
Qy 974 CTCGCTTCATCATCTGCGAGATCTGGTTTCGCCATCTCTCGATCTCTGACAGTTCC 1033
Db 882 CTCATATCTGTGATCTGTGAATTTGGTTTGCCTATGCTTGGATCTTGGACCACTTCCA 941
Qy 1034 AAGTGGTTCCTCATCGACCGGAGACGTACTCGACCCCTCTCCCTCAGGTAGCAGAGG 1093
Db 942 AAGTGGTTCCTATCGAGAGGAAACCTATCTTGACCCGCTGAGTTTAAAGGTTTGACAAG 1001
Qy 1094 GAAGGGAGCCGTCGCTGCTCGGCGGTGACCTGTTCTGTGACGACGCTGACCCGCTC 1153
Db 1002 GAAGGGCATCTTCTCAACTCGCCCTGTTGATTTCTTTGTGACGTAAGGTTGATCCCTTG 1061
Qy 1154 AAGGAGCCGCGCTGGTGACCGGCAACACCGTGTCTCTCATCTCTCGCGGTAGACTAC 1213
Db 1062 AAGGAACCTCATTGGTCACTGCTAACTGTTCTATCTATCTCTTCCGTGGAATATCCA 1121
Qy 1214 GTGACAAAGTCTCTGCTGATCTGCGACGACGCGCGGTGATGCTGACGTTTCAGATCG 1273
Db 1122 GTTGATAAGGTTTCATGCTAGCTTCTGATGATGCTGCTGACATTTGAAGCA 1181
Qy 1274 CTGTGCGAGAGCGGCTGCTGCGGCAAGTGGTGGCTTCTGCAAGAACTTGGGCATC 1333
Db 1182 TTGTCTGAACATCTGAATTTGCAAGAAATGGGTTCCTTTCTGCAAAAGATATAGCCCT 1241
Qy 1334 GAGCCCGCGCCCGGAGTCTTACTTCTCGCTCAAGGTGCACTACTCAAGCAACAAGTG 1393
Db 1242 GAGCTCGTCTCAGAGTGCTATTTCCAAAGAGATAGACTACTGTAAGAGACAGAGTG 1301
Qy 1394 CAGCCCACTTCTGTGAGAGCGCGCGCCATGAAGAGAGTATGAGAGTTCAGAGTC 1453
Db 1302 CGCCCAAACTTTGTTAGAGACGGAGAGCAATGAAGAGAGTATGAGGAATTCAGAGTC 1361
Qy 1454 CGGATCAACGCTGTTGGCCAAAGCCATGAGGTGCGGCGAGAGGCTGATCATGAG 1513
Db 1362 AGAATCAATGCTTGGTTGCTTAAAGCCCAAAAGTTCCTGAGGAAGGATGACAAATGCA 1421
Qy 1514 GACGGCACCGCTGGCCCGGAAACAAACCCGACACACCCGCGCATGATCAGGTGTTTC 1573
Db 1422 GATGGACTCATGGCCCGGAATTAATGTCCTGATCATCTGGAATGATTCAGGTTTC 1481
Qy 1574 CTGGGCCACAGCGCGGCCACGACACCGAGGGCAACGAGCTGCCCCCGCTCTGTAGCTC 1633
Db 1482 CTTGGTCAAGTGGTGGCCATGATGTGGAAGAAATGAGCTGCTCGATGTTGTTATGTT 1541
Qy 1634 TCCCGTGAGAGCGCCGGGATTCAGACACCAAGAGGCGCGGCCATGAACGCTCTG 1693
Db 1542 TCAAGAGAAAACGGCCAGGCTACAAACCATCAAGAGGCTGCTGATGAATGCAATG 1601
Qy 1694 ATTGCGCTCTCGCGCTGCTGACCAACCGCGCATTCATGCTCAATCTGAGCTGTGATCAC 1753
Db 1602 GTCCGAGTCTCTGCTGTAATTAATGCTCTTATTTGCTGAACTTGGATTTGATCAC 1661
Qy 1754 TACATCAACAGCAAGCCATCCGGAGGCCATGTGCTTCTCATGAGCCCTCAGGTC 1813
Db 1662 TATATCAATAATAGTAAAGCTATAAAGGAAGCAATGTGTTTTATGATGATCCCTTTGCTT 1721
Qy 1814 GGCCGGAAGGTCGTAGTTCAGTTCCCGCAGAGGTTGAGCGGATCGAGCTGACGAC 1873
Db 1722 GGAAGAAAGTTTGTCTATGTGCTGTTCTTCAAGATTTGATGAGGATGATCGCCATGAT 1781
Qy 1874 CGATACGCTAACAGGAACACCGTCTTCTTCACATCAACATGAAGGGCTGAGCGGATC 1933
Db 1782 CGATATGCTAACAGAAATGTTGTCTTTTTTTCGATATCAACATGAAGGTTTGGATGGTATC 1841

Qy 1934 CAAGGCCCGGTGTACGTGCGGACAGGGTGTGCTGTTCGGGCGCCAGGCGCTCTACGGCTAC 1993
Db 1942 CAGGCGCCAAATTTATGTGGTACTGTGATGTGCTTCTCAGAGCGAGGCAATTATATGGCTAC 1901
Qy 1994 AACCTCTCCCA----- 2004
Db 1902 GATGCTCCCAAAACAAAGAACCCACCATCAAGAACTTGCAACTGCTGGCCAAAGTGGTGC 1961
Qy 2005 -----GGGACCCAAAGAGCGCCAAAGATGTGTGACTGTGCGACTGTGCGCCG 2047
Db 1962 ATTTGCTGTGCTGCTGTTTGGTAAACAGGAAGACCAAGAAAGAACCAAGACCTCTAAACCT 2021
Qy 2048 TGCTTCGCGCCCAAGAG-----CGGAAACACGCGCAAGGACGGCTCGCGGAGGCG 2098
Db 2022 AAATTTGAGAAGATAAAGAACTTTTAAAGAAAAGGAAATCAAGCCCTCGCATATGCT 2081
Qy 2099 ACCGCTCATAT-----GGGAGTACATAGCAACAGGAGATGCTCATGTCC 2143
Db 2082 CTTGGTGAATTTGATGAAGCCGCTCAGAGAGCTGAAATGAAAGGCTGATTTGTAAT 2141
Qy 2144 CACATGAATTTGAGAAAGCGGTTCCGGCAGTCCGCGGGTTCGTCAAGTCGACGCTGATG 2203
Db 2142 CAACAGAAAGTTGGAAAGAAATTTGGCCAGTCTTCAGTTTTTTGTCATCCACACTTCTT 2201
Qy 2204 GAGGAAGCGCGTCCCTCTCTGTCAGCCCCCGCGGCTCTCAAGAGGAGCCATCCAT 2263
Db 2202 GAGAATGGTGAACCTTGAAGAGTGCAGTCCAGCTTCTCTCTTGAAGGAAGCTATACAT 2261
Qy 2264 GTCATCAGCTCGGCTACGAGGACAAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTAC 2323
Db 2262 GTCATCAGTTGTGATATGAAGACAAACAGCTGGGGAAGAAATATTTGGTTGGATTTAT 2321
Qy 2324 GGGTCTGATCAGGAGGACATCTCAGCGGGTTCAGATGCACTGCCGCGGTGGCGCTCC 2383
Db 2322 GGATCAGTCAAGAAATATTTCTTACTGGGTTTAAAGATGCACTGCCATGTTGGCGTCA 2381
Qy 2384 GTGTACTGATCCGAAAGCGGCGGCTTCAAGGGTTCGGGCGCGATCAATCTATCGGAC 2443
Db 2382 ATTTACTGCTATCTAAACCGGGCGCTTCAAGGTTCCGCACTCTCAATCTTTCCGAT 2441
Qy 2444 CGTCTCAACAGGTCCTCCGTTGGGCTCGTGGGCTCGTCAAGATCTTCTCAGCGCGCAC 2503
Db 2442 CGTTTTCAACAGGTTCTTCGGTGGGCTCTTGGTTCAATTTGAAATTTGTTGTCAGCAACCC 2501
Qy 2504 AGCCCCCTGTGTACGGCTACAGAACCGCAACCTCAAGTGGCTGGAGCGCTTCGCGCTAC 2563
Db 2502 TGCCCTCTCTGGTATGGTATGGTGGTGA-----CTAAAGTTCTCTGGAAAGGTTTTCTGTAC 2558
Qy 2564 ATCAACACCACTATACCCCTTCACTCGCTCCGCTGCTCGGCTACTGCAACCTCCCC 2623
Db 2559 ATTAACCTCATCGTATACCCCTTGGACATCTATCCCGCTCTTTGGGCTATTGCAACATTCGCT 2618
Qy 2624 GCGCTGCTCTCCTCAGCGGCAAGTTTCATATGCGCTCGATTAGCACTTCGCGAGATCGGTTGAGC 2683
Db 2619 GCCATCTGCTGCTGACAGGGAATTTATCAGCGCAGAGCTTAAACATGTTGCGCAGCCCTC 2678
Qy 2684 TTCTTCATCGCCCTCTTCATGTCATCTTCGCGACGGGCTCTTCGAGATCGGTTGAGC 2743
Db 2679 TGGTTTCATGTCACITTTTCATCTGCAATTTTGTCTAGAGCATCTCTGGAATGAGATGAGT 2738
Qy 2744 GGGGTGAGCATTCGAGGATGGTGAAGAACGAGCAGTTCTGGGTCTATCGGCGGCTGTTC 2803
Db 2739 GGTGTAGGCATCGATGATGCTGGTGAAGAACGAGCAGTTTGGGTCAITGGAGGCGGTGCT 2798
Qy 2804 GCGCATCTCTTCGCGCTGTCAGGGCTTCTCAAGTCTCTCGCGGATCGACCAAC 2863
Db 2799 TCACATCTCTTGTGTGTTTCCAGGACTCTCTCAAGGTCATAGCTGGTGTAGACACGAGC 2858
Qy 2864 TTACCGCTCACCTCAAGGCGCACCGGCAAGAGACGAGCTTCGCGGATCGACCAAC 2923
Db 2859 TTCACTGTGACATCCAAAG-----GGCGGAGACGAGGAGTTCTCAGAGCTGTACACA 2912

QY 2924 TTCAAGTGGACCGCTCTCATCCCGCCCAACACCGTGTCTCATATTAACGTCATCGGC 2983
Db 2913 TTCAAATGGACGACCCCTCTGTATACCTCCGACAAACCTGTCTCTACTGAACCTTCATTGGA 2972
QY 2984 GTCGTGCGCGGCATCCGAGCCATCAACACGGGTACAGTCTCTGGGGCCCTCTTC 3043
Db 2973 GTGGTAGCTGGCATCTCAATGGGATCAACACGGGATGAATCATGGGGCCCTGTTC 3032
QY 3044 GGCAAGCTCTTCTGGCTTCTGGGTCTATCGTCCACCTCTACCGTTCCTCAAGGGGCTC 3103
Db 3033 GGGAGCTCTCTTGTGATTTGGGTGATCGTCATCTTTACCGTTCCTCAAGGGTCTG 3092
QY 3104 ATGGGGCGCAGAACAGAGCCACCGTTGTGTGTCATCTGGTCCATTTCTGTGGCTCC 3163
Db 3093 GTTGGGAGGAGAACAGAGCCCAACGATTTGTCTGTGTCATCTGTGTCATCTCTGGCTTG 3152
QY 3164 ATCTTCTCCCTGCTCTGGGTTCAGGATCGACCCCTTTCATCTGTCAGGACCAAGGGCCGGAC 3223
Db 3153 ATCTTCTCGTGTCTTGGGTTCGGATCGACCCCTTCTTTCGAAAGGATGATGTCCTCCCTG 3212
QY 3224 GTCAGGCAGTGTGGCATCAATTGC 3247
Db 3213 TTGAGGAGTGTGGTCTGGATTGC 3236

RESULT 7

US-10-160-719A-57
; Sequence 57, Application US/10160719A
; Patent No. 6803498

GENERAL INFORMATION:

; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof

; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272)...(3499)

US-10-160-719A-57

Query Match

Best Local Similarity 35.1%; Score 1208.4; DB 4; Length 3704;
Matches 2077; Conservative 0; Mismatches 1046; Indels 141; Gaps 10;

QY 155 GGCACGAGACCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGTGCGAGATATGCGGC 214
Db 278 GCGCGGACGCCACGAATTCGGGGAAGCATGTGGCGGGCAGGTGTGCCAGATCTGCGGC 337
QY 215 GACGAGTGGGGTCTACGTGTGACGGCGACCTCTTTCGTCCCTCTGCAACGAGTGGGGTTC 274
Db 338 GACGGGTGGGACCGCGGGCGAGCGGACCTCTTTCACCGCTCTGCGAGTCTGCGGCTTC 397
QY 275 CCCGTGTGCGGCCCTGCTACGAGTACGAGCGCGGGAGGGGCAACGAGAACTCCCGCCAG 334
Db 398 CCCGTGTGCGGCCCTGCTACGAGTACGAGCGGCAAGGACGGCAACCCAGGGGTGCGCGAG 457
QY 335 TGCAGAGCGCTACAGCGCCCTCAAGGGGAGCCCGAGGCTTCCGGGACGATGACGAG 394
Db 458 TGCAAGACTAAGTACAGCGCCCAAGAGGGAGGCCACCACTACACGGGTGAGGAAATGAG 517

QY 395 GAGGACATCGACGACCTGGAGCACGAGTTCAACATC-----GACGACGAGAAATCAG 445
Db 518 GATGTGGATGCTGACGATGTGAGTGACTACAACTCAAGCATCTGGCAACAGGATCAG 577
QY 446 CAGAGCGAGCTGAGGGGCAACATGCAGA----ACAGCCAGATCACCGAGGCGATGTGCA 501
Db 578 AAGCAAAAGATTGCTGAGAGAAATGCTCACTTGGCGGACAAACTCAGCTGGCAGTGATATT 637
QY 502 CGGCAGGAT-----GAGCTACGGAGGGGCCCCGACGAGCGGCGGCAACNACA 551
Db 638 GGCTCGCTAAGTATGACAGCGGTGAAATTGGGCAATGGGAAGTATGACAGTGTGAGATC 697
QY 552 CCGCGCAGATCCCGGCCCATCATCACCGGCT--CCCGCTCCGTCCCGGTGAGCGGTGAGTT 609
Db 698 CTTCTGTGGATATATCCGCTCATTAATCATAGCCAGATCTCAGGAGAGATTCCTGAGGCT 757
QY 610 TCGGATTACAACGGGTATGGCCACGCGAGGTCTCGTCTTCCCTGCAAGCGCATGCCA 669
Db 758 TCCCTGATCATATGATGTCTCTGTGGGAACATTTGGCAGGCGTGACATCAATTTCTT 817
QY 670 TCCGT--ACCTGTGTCTGAGCCAGGAGTGCACAGTGGGACGAGAAGAAAGTGCAGC 727
Db 818 TATGTAATCATTTCTCCAAACCCATCGAGGAGTTCTCCGTAGACCTTGGCAATGTGCA 877
QY 728 TGAAGGAGGAGGATGACGACCTGGAAGTCCAA-----759
Db 878 TGGAAAGAGAGGTTGATGNTGGAAATAAGGTGCAATTCCTATGACCAAT 937
QY 760 -----GCAGGGCATCTCGGGCGGCGCGCGATCCCGAA 793
Db 938 GGAACAAAGCATTTGCTCCATCAGAAAGGCGTGGAGTTGCTGATATTGATGCTTCTACTGAT 997
QY 794 GACATCGAGCGCGAGCTGGCAGCTGAACGAGCGGCGAGCGCGCTGTTCGAGAAAGGTG 853
Db 998 TATAACATGGAAGATGCTTACTGAAATGATGAACCTCGGCACTCTATCTAGAAAGTG 1057
QY 854 TCGATCGCGTTCGAGCAAGGTGAACCCGTACCGGATGTTGATCGTGTGCGCTCTCGTTGTG 913
Db 1058 CCAATTCCTTCATCCAGAAATAATCGGTACAGAAATGTTGTCATGTTGGCTGTT 1117
QY 914 CTGCGCTTCTTCTCGGTACCGTATCTGCGACCCCGCTCCCGGACGCGATCGGGGTGTGG 973
Db 1118 CTATGATATTTCTTGGCTACCGTATCACACATCTCTGTGAACAAATGCATATCCCATGTGG 1177
QY 974 CTGCTCTCATCATCTGCGAGATCTGTTGCGCATCTCTGAGTCTCTCGACCAAGTTCCCC 1033
Db 1178 CTTTTATCCGTATATGTGAGATCTGGTTGCTTTGCTCTGGATTTTGGATCAGTTCCCA 1237
QY 1034 AAGTGTGTTCCCATCCACCGCGAGAGCTACTCGACCGCTCTCCCTCAGGTACGAGAGG 1093
Db 1238 AAGTGTGTTCCCATCAACCGTGAAACATACCTTGTATGAGACTGGCTTTAAGGTATGACCGA 1297
QY 1094 GAAGGGAGCGCGTGTGCTGTGTCGGCGGTGAGCTGTTGTCGTGAGACGCGGTGACCGGCTC 1153
Db 1298 GAAGGTGAACCATCTCAATTAGCTCTGTTGATATTTTGTGTCAGTACTGTGTCATCCAA 1357
QY 1154 AAGGAGCGCGCTGTGTGACCGCAACACGCTCTCCATCTCTCGCGGTAGACTACCCC 1213
Db 1358 AAGGAGCGCTCTTGTCTACTGCAATATCTGTGCTTTTCCATCTCTGCTGTCGATTTATCCG 1417
QY 1214 GTGGAACAAGGTCTCTGCTACGTTCCGACGACGCGCGCTGATGCTGTCAGTTCGAGTCG 1273
Db 1418 GTTGACAAGGTATCTTGTCTATGTTTCGGATGATGGAGCTGCTATGCTGACTTTTGTGCT 1477
QY 1274 CTGTCGAGACGCGCGAGTTCCGCGGCAAGTGGGTGCCCTTCTGCAAGAAAGTTCCGGCATC 1333
Db 1478 CTCTCTGAAACTTTAGAGTTTGTAGAAAATGGGTTCCGTTCTGTGAAGAGTCAACATA 1537
QY 1334 GAGCCCGCGCGCGAGTTCTACTTCTCGCTCAAGTTCGACTACCTCAAGGACCAAGGTG 1393
Db 1538 GAGCCTAGGGCCCCCGAATGGTACTTTGCTCAGAAAATGATTTACTTGAAGACAAAGTT 1597

1394 CAGCCACCTTCTGTCAGGCGCGCCGATGAGAGAGATGAGGATTCAGGTC 1453
1598 CAACCTCATTTGTGAAAGAACGCGCGCCATGAGAGAGATATGAGAAATTCANAATT 1657
1454 CGGATCAACGCGTGTGGCCCAAGCCATGAAGTGCAGGAGGAGGATCATCAAG 1513
1658 CGTATCAATGGTCTTTAGCCAAAGCCACAAAAGTTCCGAGGAGGATGATCATGCAA 1717
1514 GACGGCAACGCGTGGCCGCGGAAACAAACCCGCAACCCCGGCATGATCCAGGTTTC 1573
1718 GATGGTACACCTTGGCCCTGGGAAACAACTAGGAGACCATCTCGGAATGATTCAGGTTTC 1777
1574 CTGGCCACAGCGGCGCCACGACACCGAGGCGCAACGAGTGCCTCGCTGATGTC 1633
1778 CTGGGTACACAGTGGAGGCTTGACGTTGAAGGCAATGAATTCCTGTTGGTTTATGTG 1837
1634 TCCCGTGAAGACCGCCGGGATTCAGCACCAAGAGGCGCGCCATGAACGCTCTG 1693
1838 TCTCGTGAAAACGCTCTGGATTCACATCAACAGAGGCTGTGGCATGATGCACTT 1897
1694 ATTGCGTCTCCGCGTGTGACCAACGCGCATTCATGTCAACTTGGACTGTGATCAC 1753
1898 GTTCGTGTATCAGCTGCTTACTAATGGGCAATACATGTTGAATCTTGAATGAGCCAC 1957
1754 TACATCAACACAGAGCCATCCGAGGAGCCATGTCTTCATGAGACCTCAGGTC 1813
1958 TACATCAATAATAGCAAGGCTTCGAGAGCTATGTGCTTCTTATGAGACCAAACTA 2017
1814 GCGCGAAGGTCTGCTAGTTCAGTCCCGCAGAGTTTCAAGGCTGACGTCACGAC 1873
2018 GGAAGGAATGTCTGTATGTCCAAATTTCTCAGAGGTTGATGTTATGATAGGAATGAC 2077
1874 CGATACGCTAACAGGAACACGCTCTTCTCGACATCAACATGAAGGCGTGGAGCGCATC 1933
2078 CGATATGCAAAACAGGAACACGCTGTTTTTGGATTAATTAACTTGAGAGGCTTGACGGCAT 2137
1934 CAAGCCCGGTGTAGTTCGGACAGGTTGGTTCGGCGCCAGAGGCTCTACGGTAC 1993
2138 CAAGGGCCAGTTATGTGGGAACGTGGTGTGTTTAAACAGAACGGCTTATATGTTAT 2197
1994 AACCTCCCAAGGACCCAGAGGCC-----CAAGATG 2026
2198 GAGCTTCAGTCAAGAAAACAGCAGGCTTCTTCTCGCTTTGTGGGGAGGAAA 2257
2027 GTGACCTGCGACTGCTCCGCTGCTCGCGCGCAAGAGCGGAAACAGCCCAAGGAC-- 2083
2258 AAGACGTCAAAATCTAAGAGAGCTCGGAAAAGAAAGTCAACATAGACAGCGACAGT 2317
2084 -----GGGCTGCGGAGGCGCCGCTGATATGGGAGTAGAT 2119
2318 TCTGTACAGTATTTAATCTCGAAGATATAGAGGAAGGATTTGAAGGTTCTCAGTTTGTAT 2377
2120 AGCGACAAGGAGATGCTCATGTCCCATGAATTCAGAGAGCGGTTTCGGGCGAGTCCGCG 2179
2378 GATGAGAAATCGCTGATTTATGCTCAATGAGCTTGAGAGAGATTTGGCCAGTCCAGT 2437
2180 CGGTTCTCAGCTCAGCTGATGGAGAGGCGGCTGCTCTCTCGTGGAGCCCGCC 2239
2438 GTTTTGTAGCTCTACTCTGATGATATGTTGTTTCCCAATCTGCAATCTGCCAG 2497
2240 CGGCTCTCAAGAGGCGCATTCATGTATAGCTGCGGCTACGAGGACAGACGACTCG 2299
2498 TCTCTCTGAAAGAGCTATTTCAATGATCAGCTGTGGCTATGAGGACAAAACGACTCG 2557
2300 GGGCTGAGCTGGGCTGATCTACGGTTCGATCAGGAGGACATCTTCAGCGGGTTCAAG 2359
2558 GGAATGAGATTTGGTGGATCTATGGTTCTGTTACAGAGACATTCACCGGATTCAG 2617
2360 ATGCACCTGCGCGGCTGCGCTCGTGTACTGATGCGGAGGCGGGGCTTCAAGGGG 2419
2618 ATGCATCTCGAGGCTGCGCATCACTACTGATGCTTAAGCGACGAGCTTCAAGGGA 2677
2420 TCGGCGCGGATCAATCTATCGGACCGCTCTCAACAGGCTGCTCGGTTGGGCGCTGCGGTC 2479

2678 TCTGCTCCTATCAACCTTTCCGATCGTTTGAATCAAGTGTCTCGGTGGGCTCTTGGTTCC 2737
2480 GTGAGATCTTCTTACGCGGCGCAGCCCTGTGTACGGCTACAAAGAACGCAACCTC 2539
2738 ATTGMAATCTTTTACGAGGCAATGTCCTCATATGTTATGGCT---ATGGAGGCGGCTT 2794
2540 AAGTGGCTGAGGCTTTCGCTACATCAACACCAACATCTACCCCTTCACTCGCTCCG 2599
2795 AAATTCCTGAGAGATTTGCTTATCAACACAACTTTTATCCACTCACATCAATCCG 2854
2600 CTGCTGCTTACTGCAACCTCCCGGCTGCTCTCTCAGCGCAAGTTTCAATCATCGG 2659
2855 CTCCTCTGCTACTGATATTCAGGAGGTTTGTCTCTCACTGGGAAGTTTCAATCCCA 2914
2660 TCGATTAGCACTGTTCCGACGCTTCTTCTCATGCGCTCTTCAATGTCCATCTTCCGACG 2719
2915 AAGATTAGTAACCTAGAGAGTGTGTTGTTATATCGCTCTTATCTCAATCTTGGCCACT 2974
2720 GGCATCTGAGATGCGGTGAGGCGGTGAGCATTCGAGAGTGTGGAGAGACGAGAG 2779
2975 GGTATCTTGAATGAGTGTGGTGTGTCATTTGATGAATGTTGGAGAGACGAGAG 3034
2780 TTCTGGGTCTACGCGCGGTCGCGCATCTCTTCCGCTGTCGAGGCGCTCTCAAG 2839
3035 TTCTGGGTCTATGTTGTTGTTATTTTGGCGATTTATTTGCGGCTTCTTCCAGGCTCTCTGAAG 3094
2840 GTCTCGCGGATCGACACAACTTCCAGCTCACCTTCAAGGCGCCAGCGGACGAGGAC 2899
3095 GTCTGCTGATTCGACAGAGTTTCACTGTCACTTCAAGGCGCAC---TGAGAGAGAA 3151
2900 GACAGTTCCCGAGCTTACGCTTCAAGTGACACAGCTCTCTCATCTCCGCGCCACGAG 2959
3152 GGTGATTTGCGGAGCTCTACATGTTCAAGTGAGCAACGCTTCTGATCCCAACCACT 3211
2960 CTGCTCATCAATTAACGTCATCGGCGCTGCGCGGATCTCCGAGCGCATCAACAGCGG 3019
3212 ATTTGATCATCAACCTGCTGCGGTGCTGCTGCTGCTTCTTCCAGCAATCAATAGCGGT 3271
3020 TACAGTCTTGGGCGGCTTCTTCCGCAAGCTCTTCTTCCGCTTCTTGGGCTCATCGTCCAC 3079
3272 TACAGTCTGAGGAGCTCTTTTCCGGAAGCTCTTCTTGGGCTTCTGGGCTGATGTCAC 3331
3080 CTCTACCGCTTCTCAAGGCGCTCATGGGCGCGAGAACAGGAGCGCCACCGTTGTTGTC 3139
3332 CTGTACCGCTTCTCAAGGCGCTCATGGGCGGAGCAGAACCGCAGCGCCGATGTTGTT 3391
3140 ATCTGGTCCATCTGCTGGCTCCATCTTCTCTGCTGCTCTGGGCTCAGGATCGACCTTTC 3199
3392 GTCTGGGCTATCTCTCTTGGTCTGATTTTCTGCTGATGTTGGGTTCTGATTCGATTCATTC 3451
3200 ATCTGAGGACCAAGGCGCGGAGCTCAGGAGTGTGGCATCAATTTCTGAGCTGTTTAT 3259
3452 ACCACCGGCTACTGCGCTGATATCGGAATGCGGAATGCGCATCACTGCTAGGATGAGCTG 3511
3260 TAAGGTTCAAAATCTGAGCTTG 3283
3512 AAGATAGTTAAAGAGTGGAACTAG 3535

RESULT 8

US-10-160-719A-17
; Sequence 17, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A

```
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)...(3401)
US-10-160-719A-17

Query Match      35.0%; Score 1204.6; DB 4; Length 3969;
Best Local Similarity 62.8%; Pred. No. 8e-214;
Matches 2051; Conservative 0; Mismatches 1099; Indels 117; Gaps 7;

QY 89 GCGATGGAGGCCAGCGCCGGGCTGGTGGCCGGCTCGCAACCGGAAACGAGCTGGTGGCTG 148
DB 141 GCGATGGAGGCCAGCGCCGGGCTGGTGGCCGGCTCGCAACCGGAAACGAGCTGGTGGCTG 200
QY 149 ATCCGGGGCCACAGAGACCC-----AAGCCCTGGGGCGCTGAGCGGGCAGGTGTGC 202
DB 201 ATCCGGGGCCACAGAGACCC-----AAGCCCTGGGGCGCTGAGCGGGCAGGTGTGC 260
QY 203 GAGATATGGCGGACAGAGCTCGGGCTCAGGGTGGAGCGGACCTCTTCGCTCGCTGCAAC 262
DB 261 CAGATTGGCGGACAGAGCTCGGGCTTCGCCCCGGCGGGGACCCCTTCGTTGGCGGTGCAAC 320
QY 263 GAGTGGCGGCTTCCCGGCTGGTGGCCGGCTGCTACGAGTACAGCGCGGGAGGGACGCGCAG 322
DB 321 GAGTGGCGGCTTCCCGGCTGGTGGCCGGCTGCTACGAGTACAGCGCGGGAGGGACGCGCAG 380
QY 323 AACTGCCCCAGTGCAGAGCGCTCAAGCGCTCAAGGGGAGCCCGAGGGTTGCCGGG 382
DB 381 AACTGCCCCAGTGCAGAGCTCGATACAGCGCTCAAGGGGCTCAAGGGGCTCGCAACCGGT 440
QY 383 GAGGATGAGGAGGAGACATCGAGGACCTGGAGCAGAGTTCAACATCGACGACGAGAT 442
DB 441 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 500
QY 443 CAGCAGAGGCGAGCTGGAGGSCA-----ACATGACAGAACAGCCAGATCAACGAGGCGATG 496
DB 501 TCGCAGTCTGTGCCGAGTCCATGCTCTAGCGCCACATGAGCTACGCGCGTGGAGGTGAC 560
QY 497 CTGCACGGCAGGATGAGCTACGGGAGGGGCCCGGACGAGCGGCGACGCGGACAAACCCCG 556
DB 561 CCTAATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAAATGTTCCACTCTCTCAACACGGG 620
QY 557 CAGATCCCGCCCATCATCACCAGCTCCCGCTCCGCTCGCTGGCGGTGAGCGGTGAGTTCCGATT 616
DB 621 CAAATGGTGGATGACATCCACCGAGCAGCAGCGCTGGTGGCTCTTTTCATGGGTGGT 680
QY 617 ACCAAGGGGTATGGCCAGCGGAGGCTCTGCTCTTCCCTGACACAGCGCATCCATCCGTAC 676
DB 681 GGGGGAAGAGGAGTACATCCCTCTCTTATGCGGATCCCGAGTTACTGTGCAACCCAGG 740
QY 677 CTGTGTCTGAGCCAGGGAGTGCAGAGTGGGACGAGAGAAAGAGTGAAGTGGAGGAG 736
DB 741 TCTATGGACCCATCCAAAGGATCTTGTGTCATATGGGTATGGTGTGTTGCTTGGAAAGGAA 800
QY 737 AGGATCGACGACTGGAGTCCAGCAGGAGGATCTCTCGC-----GGCGGGCC 784
DB 801 CGGATGGAGAAATGGAGGACGAGACAGAGAGAGATGCAACAGCGGGGAATGATGGTGGT 860
QY 785 GATCCCGAAGACATGGAACCGGAGCTGGCACTCAACGAGCGGAGGAGCGGCTGTGTCG 844
DB 861 GGTGATGATGTCAGCATGCTGATCTACCACTTAATGGATGAGCAAGACAACTGTGCC 920
QY 845 AGGAAGGTGTGATTCGGCTCGAGCAAGGTGAACCCGTACCGGATGCTGATGCTGGTGGCT 904
```

```
DB 921 AGGAAAAATCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTAATCATTTTCG 980
QY 905 CTCGTGTGCTCGCTCTCTTCCTCGGTACCGTATCCTGCACCCCGTCCCGGACGCATC 964
DB 981 CTTGTGTTTGGGGTCTCTTCCATACCGAGTATGATCATCCGGTGAATGATGATTT 1040
QY 965 GGGCTGTGGCTCTCTCCATCATCTCGGAGATCTGGTTCGCCCATCTCTCTGGATCTCTCGAC 1024
DB 1041 GCTTTGTGGCTCATATCTGTTATCTGTGAAATCTGTTTGGCCATGCTTTGGATCTTTGAT 1100
QY 1025 CAGTTCCCAAGTGGTTCCTCCATCGACCGGAGAGCTAGCTCGACCGCTCTCCCTCAGG 1084
DB 1101 CAAATTCCTCAAGTGGTTCCTTATGAGAGAGAGACTTACCTAGACCGGCTGCTCACTGAG 1160
QY 1085 TAGGAGAGGGAAGGGAGCGCTGCTGCTGCGGGTGGAGCTGTTCTGTGAGCAGCGGTG 1144
DB 1161 TTCGACAGGAAGGCGAGCATCTCAACTTGTCTCAATGATTTCTTTGTGAGTACGGTT 1220
QY 1145 GACCCGCTCAAGGAGCGCGCTGGTGAACCGCAACACACCGTGTCTCTCAATCTCTCGCGTA 1204
DB 1221 GATCCCTTAAAGAACTCTCTTTGGTCAACAACAATACTGTTCTATCTATCTCTTCCGTTG 1280
QY 1205 GACTACCCGCTGGACAAGGTCTCTGTACGTCTCGACGACGGCGGTGCGATGCTGACG 1264
DB 1281 GATTATCTCTGTTGATAAGGTTCTTGTCTATGTTCTGTATGATGCTGCTCAATGCTAAACG 1340
QY 1265 TTCGAGTCTGCTCGGAGACGCGGAGTTCGCGCGCAAGTGGGTGCTCTTCTGCAAGAAG 1324
DB 1341 TTTGAAGCAATTAATCTGAAACATCTGAATTTGCAAGAAATGGGTCTCTTTCTGCAACCG 1400
QY 1325 TTCGGGATCGAGCCCGCGCGCGGAGTTCTACTTCTCGCTCAAGTTCGACTACCTCAAG 1384
DB 1401 TACAATATTGAACCTCGCGCTCCAGAGTGTGTACTTCCACAGAGATAGACTACTTTGAAA 1460
QY 1385 GACAGGTGTCAGCCACCTTCTGTGACGAGCGCGCGCATGAAGAGAGATGATGAGGAG 1444
DB 1461 GACAAAGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATGATGAGGAA 1520
QY 1445 TTCAGAGTCCGATCAACCGCTGTGGCCAAAGGCTCAAGGTCGCGCAGAGGGGTG 1504
DB 1521 TTCAGAGTGAAGTCAATGCTCTTAGTTGCCAAAGCCAGAAAGTCTCTGAAGAGATGG 1580
QY 1505 ATCAAGAGGACGCGACGCGTGGCCCGGGAACAACACCGCGACACCCCGCATGATC 1564
DB 1581 ACAATGCAAGATGGAAACCCCTGGCTGGAACAATGTTCTGTGATCATCTCTGGAATGAT 1640
QY 1565 CAGGTGTTCTTGGGCCACAGCGCGGCTCAGCAGACGAGGGGCAACGAGCTGCCCCGCTC 1624
DB 1641 CAGGTCTTCTTGGCCAAAGCGGAGGCTTGAATGAGGGGAAATGAACTGCCACGATTG 1700
QY 1625 GTGTAGTCTCCGTTGAGAAAGCGCGGATTCAGACCAACAAGAGCGCGGCGCATG 1684
DB 1701 GTTTATGTTCTAGAGAGAAACGACAGGCTATACCATCATAAGAAAGCTGGTGTATG 1760
QY 1685 AACGCTCTGATTTGGGCTCTCCCGCTGTGACCAACGCGCCATTCATGCTCAACTTTGGAC 1744
DB 1761 AATGCAATGGTCCGAGTCTCTGCTGACTTAAACAAATGCTCCATATTTGTTAACTTTGAT 1820
QY 1745 TGTGATCACTACATCAACAGAGGCGCATCCGGAGGCGCATGCTCTTCTCTCATGGAC 1804
DB 1821 TGTGATCACTACATCAACAGAGGCTATAAAGGAAAGCAATGTGTTTATGATGAGAC 1880
QY 1805 CCTCAGGTCCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1864
DB 1881 CTTTACTAGGAAGAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1940
QY 1865 GTGCACGACGATCGCTAACAGGAAACCGCTCTTCTTCTGACATCAACATGAAGGGGCTG 1924
DB 1941 CGCCATGACCGATATGCTAAACCGGAATGTTGCTTTTTTGTATATCAACATGAAGGTTG 2000
QY 1925 GACGGCATCAAGGGCGGCTGCTGCTGCGGACAGGGTGGTGTTCGGCGCCAGGCGCTC 1984
```

Db 2001 GATGGTATTGAGGTCCAAATTTATGTTGGTACTGATGTGTATTTAGAGCGAGGCATTAA 2060
Qy 1985 TAGGGCTAACACCTCCCAAGGGACCAAGAGGCC-----CAAGATG 2026
Db 2061 TATGGTTATGATGCCCCCAAAACAAAGAAAGCCACCATCAAGGACTTGCAACTGCTGGCCC 2120
Qy 2027 GTGACCTGCGACTGCTGCCGCTGCTTCGCGCGCAAGAGGAAAGCAACGCAAGGACGGG 2086
Db 2121 AAGTGGTCTTTTGTCTGCTTTGTCATAGGAAGCAAAAGAAAGACTTACCAAAACC 2180
Qy 2087 CTCGCGGAG----- 2095
Db 2181 AAAACAGAGAAGAAAGTTATTTTCAAGAAAGAGAACCAATCCCCTGCAAT 2240
Qy 2096 -----GGCACCGCTGATATGGAGTAGATACGCAAGAGAGATGCTCATG 2140
Db 2241 GCTCTTGGTGAATTTGAGAGCTGCTCCAGGAGCTGAGAATGAAAGGCGGTATTGTA 2300
Qy 2141 TCCACATGAATTCGAGAGAGCGTTCCGGCAGTCCCGCGGTTTCGTCAAGTCAAGCTG 2200
Db 2301 AATCAACAAAAATTTAGAAAGAAATTTGGCCAAATCTTCTGTGTTTGTATCAATCCACACT 2360
Qy 2201 ATGAGGAAGCGCGCTCCCTCTCTGTCGAGCCCGCGCTCTCAAGGAGGCCATC 2260
Db 2361 CTCGAGAAATGGTGAACCTTGAAGAGTGAAGTCTGCTTCTCTTTTGAAGAGCTATA 2420
Qy 2261 CATGTATAGCTGCGGCTACGAGGACAAGACCGACTGGGGCTGGAGCTGGGGTGGATC 2320
Db 2421 CATGTCAATTAGTTGTGTTATGAAGACAAGACAGACTGGGGAAGAGATTGGCTGGATC 2480
Qy 2321 TAGGGTCAATCAGGAGGACATCCTGACCGGGTTCAAGATGCACTCCCGGGTGGCGC 2380
Db 2481 TATGGATCAATTACAGAAAGATTTCTAACTGGTTTCAAGATGCAATGCTATGTTGGCGG 2540
Qy 2381 TCCGTGTACTGCTACCGAAGCGGGCGGCTTCAAGGGGTCCGGCGGATCAATCTATCG 2440
Db 2541 TCAATTTACTGATACCTAAACGGGTGCAATCAAGGTTCTGCACCTTGAATCTTCA 2600
Qy 2441 GACCGTCTAACACGAGTGTCCGGTGGCGCTGGGGTCCGTTCGAGATCTTCTTCAGCGG 2500
Db 2601 GATCGTCTTCCACAGGTGCTTCGGTGGCTCTTGGGTCTATTGAGATCTTCTTCAGCAAT 2660
Qy 2501 CACAGCCCCGTGTAGGCTACAGAAAGGCAACCTCAAGTGGCTGGAGCGCTTCGCC 2560
Db 2661 CATTGCCCTCTTGGTATGGGT---ATGGTGGCGGTCTGAAATTTTGGAAAGATTTTC 2717
Qy 2561 TACATCAACACCACTACCTCCCTTCACTCGCTCCCGCTGCTCGCTTACTGCACCTC 2620
Db 2718 TACATCAACTCCATCGTGTATCTTGGAGCAATCTATCCCTCTTGGCTTACTGTACATTG 2777
Qy 2621 CCCCGGTCTGCTCTCACCGGCAAGTTCAATCATGCCGTGCAATAGCAAGTTTCGCCAGC 2680
Db 2778 CCGTCCATCTGTTTATTGACAGGGAATTTATCACTCCAGAGCTGAAATATGTTGCCAGC 2837
Qy 2681 CTCTTCTTATGCGCTTCTCATGTCACTTTCGCGACGGGCAATCTCGGAGATGCGGTGG 2740
Db 2838 CTGTGGTTCTGTCACCTTTTATCTGCAATTTTGTCTACGAGCATCTCTAGAAATGAGATG 2897
Qy 2741 AGCGGGTGAAGCATCGAGGAGTGTGGAGGAAGAGCAGTCTTGGGTCTATCGCGCGGTG 2800
Db 2898 AGTGGTGTGGAATGTATGCTGTGGAGGAATGAGCAGTCTGCGGTCTATGAGAGTGTG 2957
Qy 2801 TCCGCGCATCTCTTGGCGCTGTCGAGGGGCTGTCAAGGTCTCTCGCGGGATCGACACC 2860
Db 2958 TCCTCACACCTCTTGTGTGTTCCAGGAGCTTCTCAAGGTCTAGCTGTGTGTATACA 3017
Qy 2861 AACTTCAACCTGACCTCAAGGCCACCGGCGAAGAGACAGAGTTTCGCGGAGCTCTAC 2920
Db 3018 AGCTTCAACCTGACATCAAG-----GGTGGAGATGATGAGGAGTTCTCAGAGCTATAT 3071
Qy 2921 GCCTTCAAGTGGACCAAGCTCTCATCCCGCCCAACAGCGTGTCTCATCATTAACGTCATC 2980
Db 3072 ACATTCAAATGGGACTACCTTATTGATACCTCTTACCACCTTGTCTTCTATTGAACCTTCATT 3131

Qy 2981 GGCGTGTGGCCCGGCATCTCCGACGCGCATCAACAAACGGGTACCAGTCTCTGGGGGCCCTC 3040
Db 3132 GGTGTGTGCTGCTGGCGTTTCAAAATGGCATCAATAACGGATATGAGTCAATGGGGCCCTC 3191
Qy 3041 TTCCGCAAGCTCTTCTTGGCTTCTGGTTCATGTCACCTCTACCGTTTCTTCAAGGG 3100
Db 3192 TTTGGGAGCTATCTTTGCAATTTGGTGATTTGTCATCTTTATCCCTTCTCNAAGGT 3251
Qy 3101 CTCATGGGCGCCAGAACAGGACGCCACCGTTTGTTCATCTGTGTCATTTCTGTGCC 3160
Db 3252 TTGGTTGGAAGGCAAAACAGGACCAACGATTTGTCATGCTCTGGTCCATTTCTGTGGT 3311
Qy 3161 TCCATCTTCTCCCTGCTCTGGGTGAGATCGACCTTTTCATGTCAGGACCAAGGGCCG 3220
Db 3312 TCAATCTTCTGCTCTCTTGGGTTCGGAATGATCTTCTTCCGAGGATGATGTCGCG 3371
Qy 3221 GACGTCAAGGAGTGTGSCATCAATTGC 3247
Db 3372 CTTCCTTGGAGGTGTGTGTTGGATTGC 3398

RESULT 9

US-10-160-719A-37
; Sequence 37, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (144)...(3401)
US-10-160-719A-37

Query Match 35.0%; Score 1204.6; DB 4; Length 3969;
Best Local Similarity 62.8%; Pred. No. 8e-214;
Matches 2051; Conservative 0; Mismatches 1099; Indels 117; Gaps 7;

Qy 89 GCATGAGGCGCCAGCCCGGGCTGTGGCGGCTTCGACAAACCGGAAACAGCTGGTGTG 148
Db 141 GCGATGAGGCGGAGCGCCGGCTGTGGCGGCTCCACAAACCGGAAACAGCTCGTGTG 200
Qy 149 ATCCGGGGCCACAGGACCC-----NAGCCGCTGGGGCTGAGCGGGCAGGTGTG 202
Db 201 ATCCGGCGGACGCGCATCCCGGGCGGAAAGCCCGCGGAGCAGAACCGGCGAGGTGTG 260
Qy 203 GAGATATGGCGGCGACAGAGTCCGGCTCACGGTGGAGCGGACCTCTTCGTGCTGCAAC 262
Db 261 CAGATTTGGCGCGACGACGTGCGCTTGCCCCGGGGGACCCCTTCGTGGGTGCAAC 320
Qy 263 GAGTGGCGTTCCCGTGTGCCCGCTGCTACAGTACGAGCGCGGGAGGACAGCAG 322
Db 321 GAGTGGCGTTCCCGTGTGCCCGCTGCTACGAATACGAGCGCGGGAGGACAGCAG 380
Qy 323 AACTGCCCCAGTGTGAAGAGCGGCTCAAGCGCTCAAGGGAGCCGAGGTTCGCGG 382

Db 381 AACTGCCCCAGTGCAGACTCGATACAGCGGCTCAAGGGCTGCCAACGTGTGACCGGT 440
Qy 383 GACGATGACGAGGAGGACATCGACGACTTGGAGCAGAGTTCAACATCGACGACGAGAAT 442
Db 441 GACGAGGAGGAGGACGCGTGCATGACTCTGGACAAAGAGTTCACTGGGAGGSCCATGAC 500
Qy 443 CACGAGGAGGAGTGGAGGCA-----ACATCGAGAAAGCAGAGATCACCGAGGCGATG 496
Db 501 TCGCAGTCTGTGCGCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGAGGTGAC 560
Qy 497 CTCGACGGCAGGATGACTACGGGAGGGCCCCGACGACGGCGACGGCAACACCCCG 556
Db 561 CCTAATGGCCGCCACAGCTTTCAGCTCAACCCCAATGTTCCACTCTCTCAACAACGGG 620
Qy 557 CAGATCCCGCCCATCATCACCGCTCCCGTCCCGTCCGGTGAGCGGTAGTTCCGATT 616
Db 621 CAAATGGTGATGACATCCACCGGAGCAGCAGCGCTGGTCTTTTCATGGGTGGT 680
Qy 617 ACCAAGGGTATGGCCACGGCGAGGTCGTCTTCCGTGCAAGCGCATCCATCCGTAC 676
Db 681 GGGGGAAGAGGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGG 740
Qy 677 CCGTGTCTGAGCCAGGAGTGCAGTGGGACGAGAGAAAGTGAAGTGAAGTGAAGGAG 736
Db 741 TCTATGACCCATCCAGAGATCTTGCTGCATATGGGTATGGTATGTTGCTTGGAAAGGA 800
Qy 737 AGGATGAGCAGCTGGAAGTCCAGCAGGGCATCTCGGC-----GGCGGCGCC 784
Db 801 CGGATGGAGAAATTGGAAGCAGACAGACAAGAGAGATGCACCAGAGGGGAATGATGGT 860
Qy 785 GATCCCGAAGACATGAGCGCGAGTGGCACTGAAAGACGAGCGAGGACGCGCTGTGCG 844
Db 861 GGTGATGATGGTGACGATGCTGATCTACCACTAATGGATGAGCAAGACAACTGTCC 920
Qy 845 AGGAAGGTGCGATCCGTCGACGCAAGGTGAACCCGTACCGGATGCTGATCGTGGT 904
Db 921 AGGAAATTCACCTTCCATCAAGCCAGATTAATCCATATAGGATGATTAATCATATT 980
Qy 905 CTGCTGTGTGCTCGCTTCTTCCGTACCGTATCTGCACCCCGTCCCGGACGCCATC 964
Db 981 CTGTGTGTTTGGGTTCTTCTTCCACTACCGAGTATGATCCGGTGAATGATGATTT 1040
Qy 965 GGGCTGTGCTGTCTCATCATCTCGGAGATCTGGTTGCGCATCTCTGGATTCCTCGAC 1024
Db 1041 GCTTGTGGCTCATCTGTATCTGTGAATCTGGTTTGCCATGTCTTGATTTCTTGAT 1100
Qy 1025 CAGTTCCTCCAAAGTGTTCCTCATCGACCGCGAGAGCTTACCTCGACCGCTCTCCCTCAG 1084
Db 1101 CAAATCCCAAGTGTTCCTTATGAGAGAGACTTACCTAGACCGCTGTCACTGAGG 1160
Qy 1085 TACGAGAGGAGGAGCGCTGCTGCTGCGCGGTGACCTGTTCTGTGAGCAGCGTG 1144
Db 1161 TTCGACAAAGAGGCGCGCATCTCAACTTGTCTTCCAAATTGATTTCTTGTGAGTACG 1220
Qy 1145 GACCCGTCAAGAGCGCGCTGGTGAACGGCCAAACACCGTGTCTCTCAATCTCGCGGTA 1204
Db 1221 GATCCCTTAAAGGAAACCTCTTTGGTCAACAAATACTTCTATCTATCTCTTTCGGTG 1280
Qy 1205 GACTACCCGTGACAGGTTCTCTGTAGTCTCCGACGACGCGCGGTGATGCTGACG 1264
Db 1281 GATTATCTGTGATAGGTTTCTTGCTATGTTTCTGATGATGGTGTGCAATGCTAAACG 1340
Qy 1265 TTCAGTCTGTGCGAGACGGCGAGTTTCGCGCAAGTGGGTGCGCTTCTTCCAAAGAG 1324
Db 1341 TTTGAAGCATTAATCTGAAACATCTGAATTTGCAAGAAATGGGTTCCTTCTCCAAACGG 1400
Qy 1325 TTCGGATCGAGCCCGCGCGAGTTCTACTTCTGTCTCAAGTTCAGTCTCACTCAAG 1384
Db 1401 TACAATATTGAACCTCGCGCTCCAGAGTGTACTTCCAAACAGAGATAGACTACTTGA 1460
Qy 1385 GACAAGGTGACGCCCACTTCTGTGAGGAGCGCGGCCCATGAAGAGAGATGATAGGAG 1444
Db 1461 GACAAGGTGACCAAACTTTGTAGGGAGGAGAGCAATGAAGAGAGTATGAGGA 1520

Qy 1445 TTCAAGGTCGGATCAACCGCTGTGTGGCCAAAGGCCATGAAGTGCCTGGCAGAGGGTGG 1504
Db 1521 TTCAAGGTGAGAATCAATGCCTTAGTTGCCAAAGCCAGAAAGTTCTCTGAAGAAGGATGG 1580
Qy 1505 ATCATGAAGGACGGCAGCGCTGGCCCGGGAAACAACCCGGACACACCCCGCATGATC 1564
Db 1581 ACAATGCAAGATGGAAACCCCTGGCTGGAAACAATGTTCTGTGATCATCTCGGAATGAT 1640
Qy 1565 CAGGTGTTCTCGGCCACAGCGCGCCACGACACCGAGGCAACGAGCTGCCCCGCCCTC 1624
Db 1641 CAGGTCTTCTTGGCCAAAGCGAGGCCCTTGACTGTGAGGAAATGACTGCCACCATTTG 1700
Qy 1625 GTGTACGTCTCCGTGAGAGCGCCCGGATTTCCAGCACCAACAAGAGGCGCGCGCATG 1684
Db 1701 GTTTATGTTTCTAGAGAGAAACGACCGAGCTATAACCATCATAAAGAAAGCTGGTGCTATG 1760
Qy 1685 AAGCTCTGATTCGGGTCTCCCGCTGTGACCAACGCGCCATTCATGTCTAACTTGGAC 1744
Db 1761 AATGCAATTGGTCCGAGTCTCTGTGTAATAACAAATGCTCCATATTTGTTAAACTTTGGAT 1820
Qy 1745 TGTGATCACTACATCAACAACAGCAGGCGCATCCGGGAGGCCATGTGCTTCTCTCATGGAC 1804
Db 1821 TGTGATCACTACATCAACAACAGCAGGAGGCTATAAAGGAAGCAATGTGTTTATGATGGAC 1880
Qy 1805 CCTCAGGTCCGCGGAGGTCTGTACGTTCTAGTTTCCGCGAGAGGTTTCGACGSCATCGAC 1864
Db 1881 CCTTTACTAGGAAAGAGGTTTGTGTATGTACAGTTCCCTCAAAGATTTGATGGATTGAT 1940
Qy 1865 GTGACGACGATACGCTAAACAGGAACACCGTCTTCTTCGACATCAACATGAAGGGCTG 1924
Db 1941 CGCCATGACCGATATGCTAAACCGGAATGTGTCTTTTGTATCAACATGAAGGTTTG 2000
Qy 1925 GACGGCATCAAGGCCGCGTGTACGTGCGGACAGGCTGTGTTCCGGCGCCAGGCGCTC 1984
Db 2001 GATGGTATTCAGGTTCCAAATTTATGTTGTTACTGATGTGATTTAGAAAGCAGGCAATTA 2060
Qy 1985 TACGGCTACAACCTCCAAAGGAGCCCAAGAGGCC-----CAAAGATG 2026
Db 2061 TATGGTTATGATGCCCCCAAAACAAGAAAGCCACCATCAAGGACTTGTCAACTGCTGCC 2120
Qy 2027 GTGACCTGCGACTGTGCGCCGCTTTCGCGCGGACGGAAGCGGAACAACGCCAAGGACGG 2086
Db 2121 AAGTGGTGTCTTTGCTGTGCTGCTTTGGCAATAGGAAGCAAAAGAAAGACTACCAACCC 2180
Qy 2087 CTGCGGAG----- 2095
Db 2181 AAAACAGAGAGAAAGATTTATTTTTCAGAAAGAGAGAAACAATCCCCCTGCATAT 2240
Qy 2096 -----GGCAACCGCTGATATGGAGTAGATAGCGACAAAGAGATGCTCATG 2140
Db 2241 GCTCTTGGTGAATTTGACGAAGCTGCTCCAGGAGCTGAGAAATGAAAAGGCCGGTATTGTA 2300
Qy 2141 TCCCATCATGAATTCGAGAAAGCGTTTCGGGCGAGTCCGGCGGCTTCTGTACGTCGACGCTG 2200
Db 2301 AATCAACAAAAATTAGAAAAGAAATTTGGCCAAATCTCTCTGTTTGTATCATCCACACTT 2360
Qy 2201 ATGGAGGAGGCGGCTGCTCTCTTGTGAGCCCGCGCGCTCAAGGAGGCGCATC 2260
Db 2361 CTCGAGAAATGGTGAAACCTTGAAGAGTGCAGTCTGTCTTCTTTTGAAGAGCTATTA 2420
Qy 2261 CATGTCTACAGCTGCGGCTACGAGGACAAAGACCGACTGGGGGCTGGAGCTGGGGTGGATC 2320
Db 2421 CATGTCTATTGTTGTTTATGAAGACAAAGACAGACTGGGGGAAAAGAGATTTGGCTGGATC 2480
Qy 2321 TAGGGTTCATCAGGAGGACATCTGACGGGGTTCAAGATGCACTGCCCGGGTGGCGC 2380
Db 2481 TATGGATCAGTTACAGAAATATTTAACTGGTTTCAAGATGCAATTTGTCTGTTTGGCGG 2540
Qy 2381 TCCGTGTACTGATGCCGAAGCGGGGCTTCAAGGGGTGCGCGCGGATCAATCTATCTG 2440
Db 2541 TCAATTTACTGTACCTTAAACGGGTGCAATTCAAAGGTTCTGCACCTCTGATCTTTCA 2600

2441 GACCGTCTCAACAGGCTCGCGTGGGGTCCGTCGAGATCTTCTCAGCGG 2500
Db |||||
2601 GATCGTCTTCAACAGGCTCTCGGGTCTTGGGCTATGTAGATCTTCTCAGCAAT 2660
Qy |||||
2501 CACAGCCCCCTCTGTAGGCTCAAGAAACGGAACCTCAAGTGGCTGGAGCGCTTCGCC 2560
Db |||||
2661 CATTGCCCTCTTGGTATGGGT---ATGGTGGCGGTCTGAATTTTGGAAAGATTTTCC 2717
Qy |||||
2561 TACATCAACACACCATCTACCCCTTCACTCGCTCCCGCTCGCGCTACTGCAACCTC 2620
Db |||||
2718 TACATCAACATCCATCGTATCTTGGACATCTATTCCTCTTGGCTTACTGTACATG 2777
Qy |||||
2621 CCCCGCTCTCGCTCTCACCGGAGTTCTATCTGCGTCTGATAGACCTTCGCCAGC 2680
Db |||||
2778 CTTGCCATCTGTTATTACAGGGAATTTATCACTCCAGAGCTGAATAATGTTGCCAGC 2837
Qy |||||
2681 CTCTTCTTATCGCCCTCTTCAATCTTTCGCGAGCGGCATCTTCGAGATCGGGTGG 2740
Db |||||
2838 CTGTGGTTCAATGTCATCTTTTATCTGCAATTTTGTCTACGAGCATCTTAGAATGAGATGG 2897
Qy |||||
2741 AGCGGGTGAGCATCGAGGAGTGTGGAGGAACGAGCAGTTCTGGGTCTATCGCGGGGTG 2800
Db |||||
2898 AGTGGTGTGGAATTTGATGACTGTGGAGGAATGAGCAGTTCTGGGTCAATGGAGGTGTG 2957
Qy |||||
2801 TCCGGCATCTCTTCCCGCTCTGCGAGGCTCTCAAGTCTCTCGCGGATCGACACC 2860
Db |||||
2958 TCCTCACACCTCTTGTGTGTTCCAGGGAATCTTCAAGGTCTAGCTGTGTTGATACA 3017
Qy |||||
2861 AACTTCAACCTGCACTCTCAAGCGCACCGCGAGGAGCAGCAGTTCTCGCGAGCTCTAC 2920
Db |||||
3018 AGCTTCAACCTGCACTCAAG-----GTTGGAGATGATGAGGAGTTCTCAGAGCTATAT 3071
Qy |||||
2921 GCCTTCAAGTGAGCAGCCTCTCATCCCGCCCAACCGCTGCTCATCTTAAGCTCATC 2980
Db |||||
3072 ACATTCAAATGCACTACCTTATGATACCTCTACCACTTGTCTTATTGAATTCATT 3131
Qy |||||
2981 GGCCTGTTGGCGGATCTCGAGCCCATCAACACGGTACCACTCTCGGGGGCCCCC 3040
Db |||||
3132 GGTGTGTCGTGGCGTTTCAATGCGATCAATAACGGATATGATCATGGGGCCCCC 3191
Qy |||||
3041 TTCGCAAGCTCTTCTTCCCTTCTGGGTCTATCGTCAACCTCTACCGTTCTCTCAAGGG 3100
Db |||||
3192 TTTGGGAGCTATCTTTGGATTTGGGTGATTTGTCATCTTTATCCCTTTCTCAAGGT 3251
Qy |||||
3101 CTATGGGGCCCAAGACAGACGCCACCGTTGTGTCATCTGTCATCTCTGCGCC 3160
Db |||||
3252 TTGGTTGGAGGCAAAACAGACACCAACGATTTGTCTGCTGCTGCTCAATCTCTGCTG 3311
Qy |||||
3161 TCCATCTTCTCCCTGCTCTGGTCTAGGATCGACCTTTTCATCTGTCAGGACCAAGGGCCG 3220
Db |||||
3312 TCAATCTTCTCGCTCTTGGGTTCGGATTGATCTTTCTTGGAGGATGATGGTCCG 3371
Qy |||||
3221 GACGTCAAGCAGTGTGGCATCAATTGC 3247
Db |||||
3372 CTTCTTGGAGAGTGTGTTGGATTGC 3398

RESULT 10

US-10-160-719A-13
; Sequence 13, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjarie, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (179) ... (3400)
US-10-160-719A-13

Query Match 34.5%; Score 1187.2; DB 4; Length 3725;
Best Local Similarity 62.8%; Pred. No. 1.3e-210;
Matches 2033; Conservative 0; Mismatches 1123; Indels 82; Gaps 9;

Qy 89 GCGATGGAGCGCAGCGCGCGTGGTGGCGGCTCGCACCAACCGGAACGAGCTGGTGTG 148
Db |||||
176 GAGATGGCGGCCAACAGGGGATGGTGGCAGGCTCTCACAAACCGCAACGAGTTGTCATG 235
Qy 149 ATCCGGGGCCACGAGGACCCC-----AAGCCGCTGCGGGCGCTGAGCGGGCAG 196
Db |||||
236 ATCCGCCACGAGCGCGAGCGCCCTGTCCCGCTAAGCCACGAAGAGTGGCAATGGGCGAG 295
Qy 197 GTGTGCAGATATGCGCGGAGCAGAGTGGGCTCAGCGTGGACGCGGACCTCTTCTGTCGCC 256
Db |||||
296 GTCTGCCAGATTTGTGGCGACACTGTTGGCGTTTTCAGCCACTGGTGATGTCTTTGTTGCC 355
Qy 257 TGCACACAGTGGCGCTTCCCGCTGTCGCGGCCCTGCTACGAGTACGAGCGCGGAGGGC 316
Db |||||
356 TGCATAGTGTGCTTCCCTGCTCTGCGGCCCTTGTATGATGACGAGCGCAAGAGGG 415
Qy 317 ACGCAGAACTGCCCCAGTGCAGACGCGCTTACAAGCGCCCTCAAGGGAGCGCCGAGGTT 376
Db |||||
416 AACCAATGCTGCCCTCACTGTCAGTCAAGACATAGATACAGAGACAGAAAGGTAGCCCTCGAGTT 475
Qy 377 GCGGGGAGCATGACGAGGAGGACATCGACGACTGTGAGACGAGTTCAACATCGACAC 436
Db |||||
476 CATGCTGATGATGAGGAGGAAGATGTTGATGACCTCGGCAATGAATTTCAACTATAAGCAA 535
Qy 437 GAGATTCAGCAGAGGCGAGCTGGAGGCAACATGACAGAACCCAGATCACCAGGCGCATG 496
Db |||||
536 GGCATGGGAAGGGCC--CAGAGTGGCAGCTTCAAGAGATGACGCTGATCTGTCTTCAT 593
Qy 497 CTGCAACGCGAGGATGAGCTACGGAGGGGCGCCGACGACGCGCGGCAACAAACCCCG 556
Db |||||
594 CTGCTCCCATGACCCACACCATCGGATTCACGSCCTTACAAGTGGACACAGATATCTG 653
Qy 557 CAGATCCCGCC---CATCATACCGGCTCCGCTCCGTCGCGTGGTGGAGCGGTGAGTTTCC 612
Db |||||
654 GAGAGATCTCTGATGATCTCCCTGACCGCTCATTTCTATCCGAGTCCCAACATCGAGCTATG 713
Qy 613 GATTACCAACGGGTATGCGCACCGCGAGGCTCGTCTTCCCTGCGACAAAGCGCATCCATCC 672
Db |||||
714 TTGATCCAAGCGTTCCAGTTCTCTGTGAGGATTTGGACCCCTCGAAGACTTGAATTCCT 773
Qy 673 GTACCTCTGTCTGAGCCAGGAGTGCACAGTGGGACGAGAAAGAAAGTGAAGCTG-GA 731
Db |||||
774 ATGGGCTTAATAGTGTGACTGGNAGGAAGAGTTGAGAGCTGGAGGTTAAACAGGACA 833
Qy 732 AGGAGAGGATGGAGCATGGAAGTCCAAAGCAGGGCATCTCTCGCGGGCGGCGGATCCCG 791
Db |||||
834 AAAATATGTTGCAAGTGAATTAATAATATCCAGAGGCTAGAGGAGACATGGAGGGGAGCTG 893
Qy 792 AAGACATGGAGCGCGAGCTGGCACTGACACGAGGCGAGGAGCGCGCTGTCGAGGAGG 851
Db |||||
894 GCTCAATGGAGAGATATGCAATGTTGATGATGACGCGCTACCTTTGAGCCGCAATG 953
Qy 852 TGTGATCGGCTCGAGCAAGTGAACCCGCTACCGGATGGTGTGATCGTGGTGGCTCTCGTTG 911
Db |||||
954 TGCCAAATTTCTCAACACAGCTCAACCTTTACCGGATAGTAATCATCTCTCGTCTTATCA 1013

QY	912	TGCTCGGCTTCTTCCTCCGGTACCGGTATCCTGCAACCCCGTCCCGAGCGCCATCGGGCTGT	971	Db	2094	ATGATCCTGTTTGTGCTGAAGCTGATCTGGAACCTAACTATGTTGTTAAAGAGCTGCTGTG	2153
Db	1014	TCCTGTGCTTCTTCTTCCAAATATCGTATCAGTCACTGCTGCGTAAATGCTTATGATGTT	1073	QY	2046	CGTGTCTTCGGCCGCAAGAAGCGA-----	2069
QY	972	GGCTCGTCTCCATCATCTCGAGATCTGGTTGCGCATCTCCTGATCCTCGACCGAGTTCC	1031	Db	2154	GTAGAAGGAAGAAAGAACAGAGGTATATGATAGTCAAGCCGCTATTAAGAGAA	2213
Db	1074	GGCTAGTATCTGTTAICTGTGAGTCTGGTTTGCCTTGTCTGGCTTCTAGATCAGTTCC	1133	QY	2070	-----AACACGCCAAGGACGGGTGTCGGAGGGCACCGCTGATATCGGAGTAG	2117
QY	1032	CCAAGTGGTTCCCATCGACCGGAGACGTACCTCGACCGCCTCTCCCTCAGGTACGAGA	1091	Db	2214	CAGAACTTTTCAGCTCCCATCTTTAAACATGGAAGACATCGAGGAGGCTATTAAGAGGTATG	2273
Db	1134	CAAAATGGTATCCAATCAACCGTGAGACATATCTCGACAGGCTTGCAATTGAGGTATGATA	1193	QY	2118	ATAGCACAAAGAGATGCTCATGTCCCATGNACTTCAGAAAGCGGTTCCGGCAGTCCG	2177
QY	1092	GGGAAGGGAGCGTCTGCTGTCGGCGGTGACCTGTTCTGTAGACACGGTGGACCCGC	1151	Db	2274	AGGATGAAAGGTCAGTGTCTATGTCCAGAGGAAATGGAGAAACGCTTTGGTCAGTCTC	2333
Db	1194	GAGAGGGAGAGCCATCACAGCTGGCTCCCATTTGATGTCTTTGTCAGTACAGTGGATCCAT	1253	QY	2178	CGGCGTTCGTCACTCGACGCTGATGGAAGGCGGCGTCCCTCTCTTCGTTCAGACCCCG	2237
QY	1152	TCAAGGAGCGCGCTGTGACCGCAACACCGTGTCTCCATCCTCGCCGTAGACTACC	1211	Db	2334	CAATCTTCAITGATCCACTTTATGACTCAAGGTGGCATACCACTTCNAACAAACCCAG	2393
Db	1254	TGAAGGAACCTCCACTGATCACAGCCAACTGTTTGTCCATCTTGTCTGTGGATTACC	1313	QY	2238	CGGCGTCTCAAGGAGGCCATCATGTCTACAGCTGCGGCTACGAGGACAAACCGAAT	2297
QY	1212	CCGTGGACAAAGGTCTCCTGTAGCTCTCCGACGACGGCGGTGATGCTGACCTGTGAGT	1271	Db	2394	CTTCTCTACTGAAGGAAGCTATCCATGTTATCAGCTGTGGGTACGAGGACAAACATGAAT	2453
Db	1314	CTGTGACAAAGTCTATGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGAGT	1373	QY	2298	GGGGGCTGAGCTGGGGTGGATCTACGGGTGATCTACGGAGGACATCTGACCGGGTTCA	2357
QY	1272	CGCTGTCGGAGACGGCCGAGTTCCGCGCAAGTGGGTGCCCTTCTGCAAGAGTTCCGGCA	1331	Db	2454	GGGGAAGAGATTTGGCTGGATCTATGGTTCAGTTACAGAGGATATTTCTGACTGGGTTTA	2513
Db	1374	CTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGTTCCTTTTGAAGAACACAATA	1433	QY	2358	AGATGACACTGCCGGGTGGCGCTCGGTGCTACTGCAATGCCAAGCGGGCGCTTCAAGG	2417
QY	1332	TCCAGCCCCCGCGCGAGTCTTCTTCGCTCAGGTGCGATACCTCAAGGACNAGG	1391	Db	2514	AAATGATGACGAGGCTGGCAATCAATCTACTGCAATGCCACCAACACCTTTGTTCAAGG	2573
Db	1434	TTGAACCAAGAGCTCCGAATTTTACTTTGCTCAAAAAATAGATTACTTGAAGGACAAAA	1493	QY	2418	GGTCCGCGCGGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGGCGCTGGGGT	2477
QY	1392	TGCAGCCCACTTCGTCCAGGAGCGCGCGCCATGAAGAGAGATGAGGAGTTCAAGG	1451	Db	2574	GTTCGCACCAATCAATCTTTCTGATCGTCTTAATCAGGTGCTCCGTGGGCTCTTGGGT	2633
Db	1494	TTCAACCTTCAATTTGTTAAGGAAGACGACCAATGAGAGAGATGAGAATTCABAA	1553	QY	2478	CGTTCAGATCTTCTTCAGCGGCAAGCCCTGCTGTATCGGCTACAGAAACGCAACC	2537
QY	1452	TCCGGATCAACCGCTGTGGCCAAAGCCATGAAGTGGCGGAGAGGGGTGATCATGA	1511	Db	2634	CAGTGGAAATTCGTCTAGCAGACATTTGCTATATGTTATGGCTACAAT---GGGCGAT	2690
Db	1554	TAAGAATCAATGCCCTTGTGGCCAAAGCAGAAAGTGCCCTGAAGGGGTGGACCATGG	1613	QY	2538	TCAAGTGGCTGGAGCGCTTCGCTACATCAACACCACTACCTTCACTCCCTCGCTCC	2597
QY	1512	AGGACGCGCGCTGGCCGGGACCAACACCGCGACCAACCCCGCATGATCCAGGTGT	1571	Db	2691	TGAAGCTTTTGGAGAGGCTGGCTTACATTAACACCAATGTTTATCAATCAATCTGTTTC	2750
Db	1614	CTGATGGAATCTGTGTGGCTTGGGAATAACCTTAGGGACCAATCTCGGCATGATTCAGGTGT	1673	QY	2598	CGTGTCTCCCTACTGCAACCTCCCGCGCTCTCCCTCTCCCTCACCGGCAAGTTTCATATGC	2657
QY	1572	TCTTGGGCCACAGCGCGGCCACACACCGAGGGCAACGAGCTGCCCGCCTCGTGTACG	1631	Db	2751	CGCTTATCGCTATTGTGTGCTTCCTGCTATCTGTCTTCTTACCAATAAATTTATCATTC	2810
Db	1674	TCTTGGGCCACAGTGGTGGCTTGACACTGATGGAATGAATTAACAGCTTTGCTATG	1733	QY	2658	CGTCAATTAGCAGTTCGCCAGCCTCTTCTTCAATCGCCCTCTTCAATGCTCCATCTTCGGA	2717
QY	1632	TCTCCGTGAGAGCGCCGGGATTCAGCAACACAAAGAGCGCGGCCCATGAACGCTC	1691	Db	2811	CTGAGATTAGTAATTAATGCTGGAAATGTTCTTCAATCTCTTTTGGCTCCATTTTCGCAA	2870
Db	1734	TCTCTCGTGAAGAGACCAAGGCTTTCAAGCATCACAAAGAGGCTGGTGCATGAATGCAC	1793	QY	2718	CGGCGCATCTGGAGATGCGGTGGAGCGGGTGAGCATCGAGGAGTGGTGAGGAACGAGC	2777
QY	1692	TGATTCGCTTCGCGCTGCTGACCAACCGGCCATTCATGCTCACTTGGACTGTGNTC	1751	Db	2871	CTGGTATATTGGAGCTCAGATGGAGTGGTGGCAATGGAATTTGGTGAGGAATGAGC	2930
Db	1794	TGATTCGTGTATCTGTGTGTGACAAATGGTGCCTATCTTCTCAATGTGGATTTGTGAOC	1853	QY	2778	AGTTCCTGGGTCAATCGCGCGGTTCGCGCATCTCTTCGCGCTCTGTCGAGGGGCTGTCTCA	2837
QY	1752	ACTACATCAACAAACAGCAAGCCCATCGGAGGCGCATGTGCTTCTCATGAGCCCTCAGG	1811	Db	2931	AGTTTTTGGGTATTGGTGGCACCTCTGCCCATCTCTTCGCGGTTCACAGGGTCTGCTGA	2990
Db	1854	ATTACTTCATAGCAGCAAGCTCTTAGAAGCAATGTGCTTCATGATGGATCCAGCTC	1913	QY	2838	AGTTCCTCGCGGGATTCGACCAACTTCACCTGACCTTCAAGGGCCACCGGCGACGAGG	2897
QY	1812	TCGGCGGAAGGTGTGTACTGTTCAAGTTCGCGAGAGTTTCGACGGCATCGACGTGCAAG	1871	Db	2991	AAGTGTGGGTGGATTTGATACCAACTTCACAGTTACCTCAAAGGATC---TGATGAGG	3047
Db	1914	TAGGAAGGAATACTTGTATATGTAACAATTTCCAAAGATTTGATGGCAATTGACTTGCACG	1973	QY	2898	ACGACCAAGTTTCGCGGAGCTCTACGCTTCAAGTGGACCAAGCTCTCATCCCGCCACCA	2957
QY	1872	ACGATACGCTAACAGGAACACCGTCTTCTTCGACATCAACATGAAGGGCTCGACGGCA	1931	Db	3048	ATGGCAGCTTTGCTGAGCTATATGTGTTCAAGTGGACAGTTTGTCTCATCCCTCCGACCA	3107
Db	1974	ATCGATATGCTAATAGGAACAATAGTCTTCTTGTATATCAACATGAAGGCTCTAGATGGCA	2033	QY	2958	CGTGTCTCATTAATTAACGTTCATCGGGTCTGTGGCGGCGCATCTCCGACGCCATCAACAACG	3017
QY	1932	TCAAGGCCCGGTGTACTGTGGGACAGGTGCGTGTTCGCGCGCAGCGCTCTACGCGCT	1991	Db	3108	CTGTCTCTGTCAITTAACCTGGTCGGAATGGTGGCAGGAAATTTCTGATGCAITTAACAGCG	3167
Db	2034	TTCAAGGTCCAGTCTATGTGGGAACAGATGCTGTTTCAATAGGAGGCTTCTGTATGAT	2093	QY	3018	GGTACAGTCTCTGGGGGCCCTCTTCGCGGAGCTCTTCTCGCTTCTGGGTCTATCGTCC	3077
QY	1992	ACAACCCCT-----CCCAAGGGACCCCAAGAGGCCCAAGATGGTGAACCTGCGACTGTGCCC	2045				

1614 CTGATGGAACCTGCTGGCCCTGGGAATAAACCTAGGAGCAATCTCTGGCATGATTCAGGTGT 1673
1572 TCCTGGGCCACAGCGCGGCCACAGACACCGAGGCGCAACGAGCTGCCCCGCTCGTGTACG 1631
1674 TCTTGGGGCACAGTGGTGGGCTTGACACTGATGGAAATGAATTACACAGCTCTTGTCTATG 1733
1632 TCTCCGCTGAGAGCGCCCGGGATTCAGACACACAGAGAGCGCGGCCCATGAACGCTC 1691
1734 TCTCTCGTGAAGAGACCAAGGCTTCAGCATCACAGAAAGGCTGGTGCATGAATGAATGCAC 1793
1692 TGATTGCGCTCTCCGCGCTGCTGACCAACCGGCATTCATGCTCAACTTGGAGCTGTGATC 1751
1794 TGATTGCTGATCTGCTGCTGCTGACAAATGGTGCCTATCTTCAATGTGGATGTGACC 1853
1752 ACTACATCAACAAACAGCAAGCCCATCCGGGAGGCCATGTGCTCTCTCATGGAACCTCAGG 1811
1854 ATTACTTCAATAGCAGCAAGCTCTTAGAGAGCAATGTCTTCATGATGGATCCAGCTC 1913
1812 TCGCCCGGAGGTCTGCTAGCTTCAGTTCCTCCGAGAGGTTGACGGGCATCGACGTGACG 1871
1914 TAGGAAGGAAACCTTGTATTATGTACAAATTTCCACAAAGATTTGATGGCAATTGACTTGCACG 1973
1872 ACCGATAGCTTAACAGGAACACCGTCTCTTCGACATCAACATGAAGGGGCTCGACGGCA 1931
1974 ATCGATATGCTAATAGGAACATAGTCTTCTTGATATCAACATGAAGGTCTAGATGGCA 2033
1932 TCCAAGCCCGGTGTAGCTCGGACAGGGTGCCTGTTCCGGCGCCAGCGCTCTACGGCT 1991
2034 TCCAGGCTCCAGTCTATGTGGGAACAGGATGCTGTTTCAATAGGCAGGCTTTGTATGGAT 2093
1992 ACAACCT-----CCCAAGGACCCAGAGGCCCAAGATGGTGACCTGGACTGTGCC 2045
2094 ATGATCCTGTTTGAAGTGAAGCTGATCTGGAACCTTAACATTTGTTTGAAGAGCTGCTGTG 2153
2046 CGTGTCTCGGCCGCAAGAGCGGA----- 2069
2154 GTAGAGGAGAGAGAGAGAGAGATGATATGATAGTCAAGCGCTATTATGAAGAGAA 2213
2070 -----AACACGCCAAGGACGGGCTGCGGAGGCCACCGCTGATATGGAGTAG 2117
2214 CAGAACTTTCAGCTCCCATCTTTAAACATGAGACATCGAGGAGGTATTGAAGGTTATG 2273
2118 ATAGGACAGGAGATGCTATGCTCCACATGAATCTCGAAGAGCGGTTCCGGCAGTCCG 2177
2274 AGGATGAAGGTCAGTGTCTTATGTCCAGAGGAAATTTGGAGAAACGCTTTGGTCACTCTC 2333
2178 CGGCGTTCGTCACGTCGACGCTGATGAGGAAGCGCGCTCCCTCTCTGCTCGAGCCCG 2237
2334 CAATCTTCATGATCCACCTTTATGACTCAAGTGGCATACCACCTTCAACAAACCCAG 2393
2238 CGCGCTCTCAAGGAGGCCATCCATGTCTACGCTCGGCTACGAGGACAAAGCCGACT 2297
2394 CTCTCTACTGAAGGAGCTATCCATGTTATCAGCTGTGGGTACGAGGACAAACTGAAT 2453
2298 GGGGGCTGGAGCTGGGTGATCTACGGTGCATTCAGGAGGACATCTGACCGGGTTCA 2357
2454 GGGGAAAGAGATTGGCTGATCTATGTTTCAAGTACAGAGGATATCTGACTGGGTTA 2513
2358 AGATGCACTCCCGGGTGGCGCTCGGTGACTGTCATGCGGAAGCGGGCGCTTCAAGG 2417
2514 AATATGATCAAGAGGCTGGCAATCAATCTACTGCAATGCCACAGACCTTGTTCAGG 2573
2418 GGTCCGCGCGATCAATCTATCGGACCGTCTCAACAGGTTGCTCCGGTGGCGCTGGGT 2477
2574 GTTCTGCACCAATCAATCTTCTGATCGTCTTAATCAGGTGCTCCGTTGGGCTCTTGGGT 2633
2478 CCGTCGAGATCTTCTTCAGCGGACAGCCCCCTGCTGTAGGCTCAAGAACGGCAACC 2537
2634 CAGTGGAAATCTCTGCTTAGCAGACATTTGCTCTATATGGTATGGCTACAAT---GGGCGAT 2690
2538 TCAAGTGGCTGGAGCGCTTCGCTTACATCAACACCACTCTACCCCTTCACCTCGCTCC 2597
2691 TGAAGCTTTTGGAGAGGCTGGCTTACATTAACACCATTTGTTATCCAAATCAGATCTGTTT 2750

2598 CGCTGCTCGCCTACTGCACCCCTCCCGCGCTCTGCTCTCTCACCGGCAAGTTTCATCATGC 2657
2751 CGCTTATCGCTATTGTGTGCTTCTGCTATCTGTCTTACCAATAAATTTATCATTC 2810
2658 CGTGAATTAAGCAGTTCCGCCAGCCTCTTCTCATCGCCCTCTTCATGTCCATCTTCGCGA 2717
2811 CTGAGATTAGTAATATGCTGGAAATGTTCTTCAATCTTCTTTTGGCTCCATTTTCGCAA 2870
2718 CGGGCATCTCGGAGATGCGGTGGAGCGGGTGAGCATCGAGAGTGGTGAGGAACGAGC 2777
2871 CTGCTATATTGGAGCTCAGATGGAGTGGTGGCAATTGAAGATTGGTGAGAAATGAGC 2930
2778 AGTCTTGGGTCTATCGCGCGGTGTCGCGCATCTCTTCGCGCTCTGTCGAGGGGCTGTCTCA 2837
2931 AGTTTTGGGTATTGTTGGTGACCTCTGCCCCATCTCTTCGGGTGTTCCAGGGTCTGCTGA 2990
2838 AGTCTCTCGCGGGATCGACACCACTTCACCGTCACTCCCAAGGCCACCGGCGACGAGG 2897
2991 AAGTGTGGCTGGGATTGATACCAACTTCACAGTTACCTCAAGGGCATC---TGATGAGG 3047
2898 ACAGCAGATTTCGCCGAGCTCTAGCGCTTCAAGTGGACCAAGCTCTCTCATCCGCCCAACCA 2957
3048 ATGGCGACTTTGCTGAGCTATATGTTCAAGTGGACCAAGTTGCTCATCCCTCCGACCA 3107
2958 CGCTGCTCATCAATTAACGTCTATCGCGCTCGTGGCGGCATCTCCGACGCCCATCAACAAAG 3017
3108 CTGTTCTTGTCTATTAACCTGGTCGGAATGGTGGCAGGAATTCGTATGCCATTAAACAGCG 3167
3018 GGTACAGTCTCTGGGGGCCCCCTCTTCGCAAGCTCTTCTCGCTTCTGGGTTCATCGTCC 3077
3168 GCTACCAATCTCGGGTCCGCTCTTTGSAAGCTGTTCTTCGATCTGGGTGATCCTCC 3227
3078 ACCTCTACCGTCTCTCAAGGGGCTCATGGGGGCGCAGAACAGAGAGCGCCACCGTTGTTG 3137
3228 ATCTCTACCCCTCTCTCAAGGGTCTCATGGGAGGAGGAGGAGCGCCACCAATCGTCA 3287
3138 TCATCTGTGTCATTTCTGTGGCTTCATCTTCTCTGCTCTGTGGTGAAGATCGACCCCTT 3197
3288 TCGTTTGGTCCATCTCTCTTGGTCTATCTTCTCTGCTTGGTGAAGATCGATCCTT 3347
3198 TCATCTCAGGAC---CAAGGGCGCGAGCGTCAAGCAGTGTGCGCATCAATTTGCTGAGC 3252
3348 TCATCTCCCGACACAGAAAGCTGGCGCTTGGGCGCAATGTGGTGTGAATCTGCTGATC 3405

RESULT 12

US-10-160-719A-41
; Sequence 41, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 3725
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (179)...(3400)

US-10-160-719A-41

Query Match 34.5%; Score 1187.2; DB 4; Length 3725;
Best Local Similarity 62.8%; Pred. No. 1.3e-210;
Matches 2033; Conservative 0; Mismatches 1123; Indels 82; Gaps 9;

Qy	89	GCGATGAGGCGCAGCGCGGCTGGTGGCGGCTCGACACCGGAGAGCTGGTGGT	148
Db	176	GAGATGGCGGCAACAGAGGATGGTGGCAGGCTTCACACCGCAGAGTTCGTGATG	235
Qy	149	ATCCGGGCGCACGAGGACCC-----AAGCGCTGCGGGGCTGAGCGGGCAG	196
Db	236	ATCCGCCACGAGCGGCGCGCTCCCGCTAAGCCACGAGAGTGCGAATGGCGAG	295
Qy	197	GTGTGCGAGATATGCGGCGACGAGTGGGTCTACGGTGGACGCGACCTTCTGTCGCC	256
Db	296	GTCTGCCAGATTGTGGCGACACTGTTGGCGTTTCAGCCACTGGTGATGCTTTGGTGGC	355
Qy	257	TGCAACAGTGGCGCTTCCCGTGTGCGGCGCTCTACGAGTACGAGCGCGGAGGGC	316
Db	356	TGCAATGAGTGTGCTTCCCTGTCTGCGCGCTTGTCTATGATGATGAGTACGAGCGCAGGAGGG	415
Qy	317	ACGCAGAACTGCCCCCGCTGCAAGACGCGCTTACAGGCGCTCAAGGGGAGCGCCAGGGTT	376
Db	416	AACCAATGCTGCTCTAGTGCAGACTAGATACAGAGACAGAAAGTGGCTCGAGTT	475
Qy	377	GCCGGGACGATGACGAGGAGCATCGACGCTGGAGCAGAGTTCACATCGACAC	436
Db	476	CATGGTATGATGAGGAGGAGATGTTGATGACCTGGACATGAATCAACTATAAGCAA	535
Qy	437	GAGAAATCAGCAGGCGCTGGAGGCGCAATGACGAGACACCCAGATCACCGAGCGCATG	496
Db	536	GGCAATGGGAAGGGCC--CAGAGTGGCAGCTTCAAGGAGATGACGCTGATCTGTTCAT	593
Qy	497	CTGCACGCGCAGGATGAGCTACGGAGGGCGCCGACGACGCGGCAACACACCCCG	556
Db	594	CTGCTGCCATGACCCACACCATCGGATTCAACGCTTACAGTGGACACAGATATCTG	653
Qy	557	CAGATCCCGCC-----CATCATCACCGGCTCCCGTCCGTCGCGTGGAGCGGTGAGTTCC	612
Db	654	GAGAGATCCCTGATGATCCCTGACCGTCAATCTATCCGCGATCCAAATCGAGCTATG	713
Qy	613	GATTACCAACGGGTATGGCACGCGGAGGTCTGCTCTTCCCTGACACGACGCTCCATCC	672
Db	714	TTGATCAACGCGTTCCAGTCTCTGTGAGGATTGTGGACCCCTCGAAGGACTTGAATTCCT	773
Qy	673	GTACCTGTGTCTGAGCGAGGAGTGCCAGTGGGACGAGAAAGAGTGGAGCTGA	731
Db	774	ATGGGCTTAATAGTGTGACTGGAGGAAAGAGTGGAGCTGGAGGTTAAACAGGACA	833
Qy	732	AGGAGAGGATGAGCAGCTGGAAGTCCAAAGCGGCGATCTCCGCGCGCGCGCGATCCCG	791
Db	834	AAAATATGTTGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGACATGGAGGGACTG	893
Qy	792	AAGACATGGAGCGCGAGTGCGACTGAACGACGAGGCGAGCGACCGCTGTGAGGAGG	851
Db	894	GCTCAAAATGGAGAGATATGCAAAATGGTTGATGATGACGCGCTACCTTTGAGCGCGATG	953
Qy	852	TGTCGATCGGTGAGCAAGGTGAACCGTACCAGTGGTGGTGGTGGTGGTGGTGGTGGTGG	911
Db	954	TGCCAATTTCTCAACAGCTCAACCTTTACCGGATAGTAAATCAATCTCCGTCTTATCA	1013
Qy	912	TGCTCGGCTTCTTCTCGGTACCGTATCTGACACCCCGTCCCGGAGCGCATCCGGCTGT	971
Db	1014	TCCTGTGCTTCTTCTTCAATATCGTATCGATCATCCAGTGGTAAATGCTTTATGATGT	1073
Qy	972	GGCTCGTCTCATCATCTGCGAGATCTGGTTGGCCATCTCTGGATCCTCGACAGTTCC	1031
Db	1074	GGCTAGTATCTGTTATCTGTGAGGTCTGGTTTGGCTTGTCTCGGCTCTAGATCAGTTCC	1133
Qy	1032	CCAAGTGTTCCTCATCGACCGCAGAGTACCTCGACCGCTCTCCCTCAGGTACGAGA	1091
Db	1134	CAAAATGGTATCCAAATCAACCGTGAGACATATCTCGACAGGCTTGCATTTGAGGTATGATA	1193

Qy	1092	GGGAAGGGAGCGCTGCTGCTGCTGGCGGTGGACCTGTTGCTGAGCACGCTGACCCGC	1151
Db	1194	GAGAGGAGAGCCATCAAGCTGGCTCCCATTTGATGTTGTCAGTACAGTGGATCCAT	1253
Qy	1152	TCAGGAGCGCGCGCTGGTGAACCCCAACCGTGTCTCCATCTCGCCGTAGACTACC	1211
Db	1254	TGAAGGAACCTCCACTGATCACGCCCAACTGTTTGTCTGTCGTCGTGATGATG	1313
Qy	1212	CCGTGGCAAGGTCTCTGCTAGCTCTCCGACGCGCGGTGATGCTGATGCTGAGT	1271
Db	1314	CTGTTGACAAAGTGTCTATGTTTCTGATGATGCTCAGCTATGCTGATTTTGGT	1373
Qy	1272	CGCTGTCGAGACGCGCGGAGTTCCGCGCAAGTGGGTGCCCTCTGCAAGAGTTCGCA	1331
Db	1374	CTCTCTCTGAACTGCGCGAATTTGCTAGAAAGTGGGTTCCTTTTGAAGAGCACATA	1433
Qy	1332	TCGAGCCCCCGCGCGGAGTTCTACTTCTCGCTCAAGGTGCTACTCTCAAGGACAAAG	1391
Db	1434	TTGAACCAAGAGCTCCAGAAATTTTACTTTGCTCAAAAATAAGATTACTTGAAGGACAAA	1493
Qy	1392	TGCAGCCACCTTCGTGACGAGCGCGCCATGAAGAGAGATATGAGGAGTTCAAGG	1451
Db	1494	TTCAACCTTCATTTGTTAAGGAAGACGAGCAATGAAGAGAGATATGAAGAAATTCAAA	1553
Qy	1452	TCCGATCAACGCGCTGGTGGCCCAAGCCATGAGGTGCCGCGAGAGGGGTGATCATGA	1511
Db	1554	TAAAGATCAATGCTTGTTCCTAAAGCACAGAAAGTCCCTGAAGAGGGGTGGACCATGG	1613
Qy	1512	AGGAGCGACGCGTGGCGCGGGAACAAACCCGCGACCAACCCGCGCATGATCCAGGTGT	1571
Db	1614	CTGATGGAATCTGCTTGGCTGGGAATAACCTTAGGACCATCTCTGGCATGATTCAGGTGT	1673
Qy	1572	TCCTGGGCCACAGCGCGGCGCACACACCGAGGGCAACGAGCTGCCCGCTCGTGTACG	1631
Db	1674	TCCTGGGGCACAGTGGTGGCTTGACACTGATGGAATGAATACCACGCTCTTGTCTATG	1733
Qy	1632	TCTCCGTGAGAGCGCGCGGGATTCCAGACACCAAGAGAGCGCGCGCATGAAACGCTC	1691
Db	1734	TCTCTCGTAAAGAGACCGAGCTTTCAGCATCAACAAGAGGCTGCTGCAATGAATGCAC	1793
Qy	1692	TGATTCGCTCTCCGCGCTGCTGACCAACCGCCATTCTGCTCAACTTGGACTGTGATC	1751
Db	1794	TGATTCGTGATCTGCTGCTGACAAATGCTGCTATCTCTCAATGTGATTTGAC	1853
Qy	1752	ACTACATCAACACAGCAAGCGCCATCCGAGGAGCCATGTGCTTCTCATGAGCCCTCAGG	1811
Db	1854	ATTACTTCAATAGCAGCAAGCTCTTAGAGAGCAATGTGCTTCTCATGATGATCCAGCTC	1913
Qy	1812	TCGCGCGGAAGGTCTGCTAGTTTCACTTCCGCGAGAGTTTCGACGGCATTCGACGTGACG	1871
Db	1914	TAGGAAGGAAAACTTGTATGTAACAATTTCCACAAAGATTTGATGGCATTCGACTTGCACG	1973
Qy	1872	ACCGATACGCTTACAGGAACACCGCTCTTCTCGACATCAACATGAAGGGCTGACGCGCA	1931
Db	1974	ATCGATATGCTAATAGGAACATAGTCTTCTTTGATATCAACATGAAGGCTTAGATGGCA	2033
Qy	1932	TCCAAGCGCGGTGTACTCTGGGACAGGCTGCTGTTTCCGCGCGCACGCGCTCTACGCTC	1991
Db	2034	TTGAGGTTCAGTCTATGTGGGAACAGATGCTGTTTCAATAGGACAGGCTTTGTATGAT	2093
Qy	1992	ACAAACCT-----CCCAAGGAGCCCAAGAGGCGCAAGATGTTGACCTCGGAGTGTGCC	2045
Db	2094	ATGATCTGTTTGTGACTGAAAGTGTCTGGAACCTAAACATTTGTTGTTAAGAGCTGCTGTG	2153
Qy	2046	CGTCTTCGCGCGCAAGCGGA-----	2069
Db	2154	GTAGAAGGAAGAAAGAAACAAGAGTTATATGGAATGATCAAAAGCCGCTATATGAAGAA	2213
Qy	2070	-----AACACGCCAAGGACGGGCTCGCGAGGGGCAACGCTGATATGGGAGTAG	2117
Db	2214	CAGAACTTTCAGCTCCCATCTTTAAACATGGAAGACATCGAGAGGATTCAGAGTTATG	2273

QY 2118 ATAGCGACAGGAGATGCTCATGTCCCAATGAATTCGAGAAAGCGGTTCGGGCGATCCG 2177
DB AGGATGAAGAGTCAAGTCTTATGCTCCAGAGGAAATTGGAGAAACGCTTTGGTCACTCTC 2333
QY 2178 CGCGGTTCTGTCAGTCCGATGATGAGAAAGCGCGCTCCCTCGTTCGAGCCCG 2237
DB CAATCTTCAATGATCCACCTTTATGACTCAAGGTGGCATACCACCTTCAACAAACCCAG 2393
QY 2238 CGCGCTCTCTCAAGGAGCCATCCATGTCTCATAGCTGCGGCTACGAGGACAAACCGACT 2297
DB CTCTCTCTACTGAAGGAGCTATCCATGTTATCAGCTGTGGGTACGAGGACAAACTGAAT 2453
QY 2298 GGGGGCTGAGCTGGGGTGATCTACGGGTGATGATCGAGGAGCATCTCGACGGGTTCA 2357
DB GGGGAAAGAGATTGGCTGATCTATGGTTCACTTACAGAGGATATTCTGACTGGGTTTA 2513
QY 2358 AGATGCACTGCGCGGCTGCGCTCGCTGATGCAATGCCGAGCGGGCGGCTCAAGS 2417
DB AATGATGCAAGAGCTGGCAATCAATCTACTGCAATGCCACCAACCTTGTTCAGG 2573
QY 2418 GGTGCGGCGGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGGTGGGCGCTGGGT 2477
DB GTTCTGACCAATCAATCTTCTGATCGTCTTAATCAGGTGCTCCGTGGGCTCTGGGT 2633
QY 2478 CGGTGAGATCTTCTACGGCGGACAGCCCCCTGTGTACGGCTCAAGAAACCGCAAC 2537
DB CAGTGAATTTCTGCTTAGCAGACATTTGCTATATGATGCTACAAAT---GGGCGAT 2690
QY 2538 TCAAGTGGCTGAGCGCTCGCCCTACATCAACACCACTTACCCCTTCACTCGCTCC 2597
DB TGAAGCTTTGGAGAGCTGGCTTTACATTAACACCAATGTTTATCAATCACATCTGTT 2750
QY 2598 CGCTGCTCGCTACTGCAACCTCCCGCGCTCTGCTCTCCTCACCGGCAAGTTCAATGC 2657
DB CGCTTATCGCTATTGTGCTTCTGCTCTCTTCTTACCAATAATTTATCATTC 2810
QY 2658 CGTGCATAGCAGTTGCGCAGCTCTTCTTCAATCGCCCTTTCATGTCACTTCGGA 2717
DB CTGAGATTAAGTAATATGCTGGAATGTTCTTCACTTCTTTTGTGCTCCATTTTCGCA 2870
QY 2718 CGGGCATCTGGAGATGCGGTGAGCGGGTGAGCATCGAGAGTGGTGAGGAAACGAC 2777
DB CTGATATTTGGAGCTCAGATGAGTGGTGTGGCAATGGAATGGTGAGAAATGAGC 2930
QY 2778 AGTCTGCGTCAATCGCGGCTGCTCGCGCATCTTTCGCGCTGCTGCGAGGCGCTCTCA 2837
DB AGTTTTGGGTTATTGGTGCCACCTCTGCCCATCTCTTCGGGTGTTCCAGGCTCTCTGA 2990
QY 2838 AGTCTCTCGCGGGATGACACCAACTTCAACCTGACCTCAAGGCCACCGGCGACGAG 2897
DB AAGTGTGGCTGGGATTTGATACCACTTCAAGTTACCTCAAGGCAATC---TGATGAG 3047
QY 2898 ACAGCAGTTTCGCGAGCTCTAGCCTTCAAGTGGACCAAGCTTCACTCCCGCCACCA 2957
DB ATGGCACTTTGCTGAGCTAATATGTTTCAAGTGGACCAAGTTGCTCATCTCCCTCCGACA 3107
QY 2958 CGCTGCTCATATTAAAGTCAATCGGGCTGCTGCGCGCATCTCCGACGCGCATCAACAA 3017
DB CTGTTCTTGCTATTAACTGGTCCGAATGGTGGCAGGAAATTCGATGCAATTAACAGC 3167
QY 3018 GGTACAGTCTCGGGGCGCTCTTCGGCAAGCTCTTCTTCGCTTCGCGCTTCGCGTCA 3077
DB GCTACCAATCTCGGGGTCGCTCTTTGGAAGCTGTTCTTCTCGATCTGGGTGATCCTCC 3227
QY 3078 ACCTTACCGTTCTTCAAGGGCTCATGCGGCGCGAGACAGAGCGCCACCGTTGTTG 3137
DB ATCTTACCCCTTCTCAAGGGTCTCATGCGGAGGAGCAACCGCAACCAACAAATCGTCA 3287
QY 3138 TCACTGGTCCATTTCTGCTGGCTCCATCTTCTTCTGCTGCTGCTGGGTCAAGCCCTT 3197
DB TCGTTTGGTCCATCTCTGCTGCTATCTTCTTCTTCTGCTGCTGAGATCGATCTT 3347
QY 3198 TCATCGTAGGAC---CAAGGGCGCGGACGTACAGGAGATGTGGCATCAATTTGCTGAGC 3252

DB 3348 TCATCTCCCGACACAGAAAGCTGCGCCTTGGGGCAATGTGTGTGAACCTGCTGATC 3405
RESULT 13
US-10-160-719A-33
; Sequence 33, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 3753
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)... (3408)
US-10-160-719A-33
Query Match 34.3%; Score 1180.2; DB 4; Length 3753;
Best Local Similarity 63.0%; Pred. No. 2.6e-209;
Matches 2043; Conservative 0; Mismatches 1113; Indels 85; Gaps 11;
QY 89 GCGATGAGGCGCAGCGCGGCTGTGGCGGCTCGCACAAACCGGAACGAGCTGTGCTG 148
DB 181 GAGATGGCGCCAAACAAGGGGATGTGGCGGCTCGCACAAACCGGAACGAGTTCGTATG 240
QY 149 ATCCGGGGCCACGAGGACC-----CCAGCCGCTGCGGGCGCTGAGCGGGGCGAG 196
DB 241 ATCCGCGACAGCGCGGATGTGCGGGCTCGGCTAAGCCCAAAAGAGTGCBAATGGACAG 300
QY 197 GTGTGCGAGATATGCGCGACGAGGTGCGGCTGACGCGGACCTCTTCGTCGCC 256
DB 301 GTCTGCCAATTTGCGGTGACTCTGTGGGTGTTTCAGCCACTGCTGTCTTGTGTC 360
QY 257 TGCAACGAGTGTGCGGTTCCCGTGTGCGGCGCTGCTACGAGTACGAGCGCGGGAGGC 316
DB 361 TGCAATGAGTGTGCGTTCCTTCTGTCGCGCCCATGCTATGAGTATGAGCGCAAGGAGGG 420
QY 317 AGCAGAACTGCCCCCAGTGCAGAGCGCTCAAGCGCTCAAGGGGAGCCCGAGGTT 376
DB 421 AAACCAATGTGCCCCAGTGCAGACTAGATACAAGAGACAGAAAGGTAGCCCTCGAGTT 480
QY 377 GCGGGGAGCATGACGAGGAGGACATCGAGCACTGGAGCAAGAGTTCACATCGACGAC 436
DB 481 CATGTGATGAGATGAGGAAGATGTTGATGACCTTAGACATGATTTCAACTACAAGCAA 540
QY 437 GAGAATCAGCAGAGGAGCTGGAGGGAACAATGAGAGCAAGCCAGATCAACCGAGGCGATG 496
DB 541 GGCAGTGGGAAAGGCC--CAGAGTGGCAACTGCAAGGAGATGATGCTGCTCTTCTCAT 598
QY 497 CTGACGCGGAGATGAGCTACGGAGGGGCGCGAGCGGCGACGGCAACCAACACCCCG 556
DB 599 CTGCTCGCATGAGCCACATCATCGATTCGATTCGCGCTTCAAGCGGTCAACAGATATCTG 658
QY 557 CAGATCCCGCCCATCATCACCGGCTCCCGCTCCGT---GCCGCTCAGGCGGTGAGTTTCC 612
DB 659 GAGAGATTCCTGATGTTCCCTGACCGCTCATTTCTATTCGAGTCCCAACATCGAGCTATG 718


```
Db 2876 CCACTGGTATATTGGAGCTTAGATGGAGTGGTGTGGCATTTGAAGATTGGTGAGAAATG 2935
QY 2775 AGCAGTTCTGGGTATCGCGGGGTGTCCGCGCATCTCTTCGCGGTGTCGACAGGGCTGC 2834
Db 2936 AGCAGTTTGGGTATTGGTGGACCTCTGCCCATCTCTTCGAGTGTTCACAGGCTGC 2995
QY 2835 TCAAGTCTCGCGGGATCGACACCACTTCAACCGTCACCTCAAGGCCACCGGGACG 2894
Db 2996 TGAAGTGTGGCTGGGATTGATACCAACTTCACAGTTACCTCAAGGCATC---TGATG 3052
QY 2895 AGCAGCAGGATTCGCGGAGCTCTAGCCCTTCAAGTGGACACGCTCCTCATCCGCCCA 2954
Db 3053 AGATGGCAGCTTTCGTAGCTATATGTGTCAAGTGGACCAAGTTGCTCAATTCCTCCGA 3112
QY 2955 CCACGCTGCTCATCTTAACGTCATCGGCGCTCGTGGCGGCGATCTCCGACGCCATCAACA 3014
Db 3113 CCACGTTCTTGTTCATTAACCTGGTCGAATGTGGCAGAAATTTCTTATGCCATTAACA 3172
QY 3015 ACGGGTACAGTCTCGGGGCCCTCTTCGGCAAGCTTCTTCGCTTCGCTTCGGGTCAATCG 3074
Db 3173 GTGGCTACCAATCTCGGGTCCGCTCTTTGGAAAGCTGTCTTCGATCTGGGTGATCC 3232
QY 3075 TCCACCTCTACCGTTCCTCAAGGGGTCTATGGGGCCGAGACAGACAGCCCAACCGTTG 3134
Db 3233 TCCATCTCTACCCCTTCTCAAGGGTCTCATGGAAAGGCGAGAACCCGACACCAACATCG 3292
QY 3135 TTGTCATCTGGTCCATCTCTGCGCTCCATCTTCTCCCTGCTCTGGTCAGGATCGACC 3194
Db 3293 TCATTGCTGTCATCTCTTGGCATCTATCTTCTCTCTGCTGTGGTGAAGTCTGATC 3352
QY 3195 CTTTCATCTGTCAGAC---CAAGGGCCGAGCTCAGGCGAGTGTGGATCAATTTGCTGAG 3251
Db 3353 CTTTCATCTCCGACACAGAAAGCTGCTGCTTGGGGCAATGTGGCTCAACTGCTGAT 3412
QY 3252 C 3252
Db 3413 C 3413

RESULT 14
US-10-160-719A-53
; Sequence 53, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 2002-06-03
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 3753
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)...(3408)
US-10-160-719A-53

Query Match 34.3%; Score 1180.2; DB 4; Length 3753;
Best Local Similarity 63.0%; Pred. No. 2.6e-209;
Matches 2043; Conservative 0; Mismatches 1113; Indels 85; Gaps 11;
QY 89 GCGATGAGCCAGCGCGGGTGGTGGCCGGCTCGCACAAACCGGAACGAGCTGGTGCTG 148
```

```
Db 181 GAGATGCGCGCCCAACAAAGGGGATGGTGGCGGGCTCGCAACAACCGCAACGAGTTCTGTCATG 240
QY 149 ATCCGGGGCCACAGGAGAC-----CCAAGCCGCTGCGGGCTGAGCGGCGAG 196
Db 241 ATCCGCCACGACGGCGATGTGCGGGCTCGGCTAAGCCCAAAAAGAGTGGGAATGACAG 300
QY 197 GTGTGCGAGATATGCGGCGACGAGGTGCGGCTCACGGTGAGACGCGGACCTCTTCGTGCGC 256
Db 301 GTCTGCCAGATTTCGCGTGACTCTGTGGGTGTTTACGCCACTGGTGATGTCTTTGTGTCC 360
QY 257 TGAACAGTGTGGCTTCCCGGTGTCGCGCTCTGACAGTACGAGTACGAGGCCCGGAGGGC 316
Db 361 TGAATAGTGTGCTTCTCTGTCGCGCCATGCTATGAGTATGAGCTATGAGCGCAAGGAGGG 420
QY 317 ACGCAGAACTGCCCGCCAGTGCAGACGCGCTCAAGCGCTCAAGGGAGCCCGAGGGTT 376
Db 421 AACCAATGTGCCCCCAGTGCAGACTAGATACAAGAGACAGAAAGGTAGCCCTCGAGTT 480
QY 377 GCGGGGACGATGACGAGGAGGACATCGACGACTCGAGCAGCAGCTGAGCAGAGTTCAACATCGACGAC 436
Db 481 CATGGTGTAGGATGAGGATGAGGAGATGTTGATGACTAGACAAATGAATTTCAACTACAGCAA 540
QY 437 GAGAATCAGAGGAGGAGCTGGAGGGCAACATGCAAGACAGCCAGATCACCGAGGCGATG 496
Db 541 GGCAGTGGGAAAAGGCC--CAGAGTGGCAACTGCAAGGAGATGATGCTGATCTGTCTTCAT 598
QY 497 CTGACGCGCAGATGAGTACGGGAGGGGCGCGACGCGGACGCGCAACACACACCCCG 556
Db 599 CTGCTCGCCATGAGCCACATCATCGGATTCACGCTGACAGCGCTGCAAGCGGTCAACAGATATCTG 658
QY 557 CAGATCCCGCCCATCATCACCGGCTCCCGCTCCGT-----GCCGGTGAAGCGGTGAGTTTCC 612
Db 659 GAGAGATTCCTGATGCTTCCCTGACCGTCAATTCATCCGAGTCCCAACATCAGACTATG 718
QY 613 GATTACCAACGGGTATGGCCACGCGAGAGTCTGCTTCTCTGCAACAGGCGATTCATCC 672
Db 719 TTGATCCAAAGCGTCCCAAGTTCCTGTGAGGATTTGTGGACCCCTCGAAGCGCTTGAATTCCT 778
QY 673 GTACCTGTGTCTGACGCGAGGAGTCCCAAGTGGGACGAGAGAAAG---AAGTGAGCT 728
Db 779 ATGGGCTTAATAGTGTGACTGGAAGAAAGATTGAGAGCTGGAGGGTTAAACAGACA 838
QY 729 GGAAGGAGGAGTGGACGACTGGAAGTCCAAGCAGGGCATCTCTCGCGGGCGGCGCGATC 788
Db 839 AAAATATGATGCAAGTACTAATAATATCCAGAGCTAGAGGAGGAGACATCGAGGGGA 898
QY 789 CCGAAGACATGAGCGCGGACGTGGGCACTGAACGACGAGGGAGGAGCGCTCTGTCAGGA 848
Db 899 CTGGCTCAAAATGAGAAAGATATGCAAAATGGTGTGATGATGACGCGCTACCTTTGAGCCGTA 958
QY 849 AGGTGTGATCGGTGCGACGAGGTGAACCCGTACCGGATGGTGATCGTGGTGGTCTCG 908
Db 959 TCGTGCAAATTTCTCCAAACCGAGCTCAACCTTTACCGGGTAGTGATCAATTCCTCGCTTA 1018
QY 909 TTGTGCTCGCTTCTTCTCCGTAACCGTACCGTATCTGACACCCCGTCCGCGAGCCATCGGGC 968
Db 1019 TCATCTGTGCTTCTTCTCCAGTATCGTGTGATCATCCAGTGGCGTGTATGAT 1078
QY 969 TGTGGTCTGCTTCATCATCTGCGAGATCTGGTTCGCCATCTCTCGATCTCTCGACGAT 1028
Db 1079 TATGGCTAGTATCTGTATCTGCGAGGTCTGGTTTGCCTTGTCTTGGCTTCTAGATCAGT 1138
QY 1029 TCCCCAAGTGGTTCCCATCGACGAGAGCTGCTGCGCGGTGAGACCTGTTCTGAGCACCGTGGACC 1088
Db 1139 TCCCAAAATGGTATCAATCAACCGGAGACATATCTTGACAGGCTTGCAATGAGGTATG 1198
QY 1089 AGAGGGAAGGGAGCGCTGCTGCTGTCGCGGTGAGACCTGTTCTGAGCACCGTGGACC 1148
Db 1199 ATAGAGGGGAGAGCCATCACAGTGGTCCCAATTCATGCTGCTTCGTGAGTACAGTGGATC 1258
QY 1149 CGTCAAGAGCGCGCGCTGGTGACCGCAACCGGTCTCTCCATCTCTCGCGGTAGACT 1208
```


Db 1259 CATTGAAGGAACCTCCACTGATCAGCAACACATGTTTTTGTCTCATCTTCTTCTGTGGATT 1318
Qy 1209 ACCCGTGGACAAGGTCTCTGCTA CGTCTCGAGCAGCGCGCGTGCATGCTGACGTTTCG 1268
Db 1319 ACCCTGTGTAACAAGTGTCTATGCTATGTTTCTGATGATGGTTTCAGCTATGCTGACTTTTG 1378
Qy 1269 AGTCGTGTGAGAGCGGCCAGTTCCGCGCAGTGGTGGTCCCTTCTGCAAGAGTTTCG 1328
Db 1379 AGTCTCTCTCAGAAACCGCAAAATTTGCTAGAAAGTGGTTTCCCTTTTGTAAAGAGCAC 1438
Qy 1329 GCATCGAGCCCCCGCGCCGAGTTTACTTTCTCGCTCAAGGTGCACTACCTCAAGGACA 1388
Db 1439 ATATTGAACCAAGAGCTCCAGAAATTTACTTTGCTCAAAAATAGATTACCTGAAGACA 1498
Qy 1389 AGTGCAGCCCACTTCGTGAGAGCGCCGCGCATGAAGAGAGATGATGAGAGTTCA 1448
Db 1499 AAATCAACCTTCAATTTGTTAAGGAAGACGCGCAATGAAGAGGGAGTATGAAGAAATCA 1558
Qy 1449 AGTCCGGATCAACGCGTGTGTGGCCRAAGGCCATGAAGTCCCGCAGAGGGGTGGATCA 1508
Db 1559 AAGTAAGAAATCAATGCCCTTTTCCAAAGCAGCAAAAGTGCCTGAAGAGGGGTGGACCA 1618
Qy 1509 TGAAGGACGGCAGCGGTGGCCCGGGAACAACACCCGCGACACCCCGGCATGATCCAGG 1568
Db 1619 TGGCTGATGNAACTGCACTGGCTTGGGAATAATCTAGGGACCATCTTGGCAATGATTCAGG 1678
Qy 1569 TGTTCCTGGGCCACAGCGCGCGCCACGACACCGAGGGCAACGAGCTGCCCGCCCTCGTGT 1628
Db 1679 TTTTCTTGGGGCACAGTGGTGGGCTCGACACTGATGGAATGAGTTACCACTCTTGTCT 1738
Qy 1629 ACGTCTCCGTGAGAAGGGCCCGGATTCGAGCACCAAGAGGGCCGCGGCCATGAACG 1688
Db 1739 ATGTCTCTCGTGAAGAGACCGGGCTTTTCAAGCATCAAGAGAGGGCTGGTCAATGAATG 1798
Qy 1689 CTCTGATTCCGCTCTCGCCGTGTGACCAACGCGCATTCATGCTCAACTTGAAGTGTG 1748
Db 1799 CGCTGATTCGTGATCTGCTGTGCTGACAAATGGTGCCTATCTTCTCAATGGATTTGCG 1858
Qy 1749 ATCACTACATCAACAACAGAGCCCATCCGGAGGGCCATGTGTTCTCTCATAGACCCCTC 1808
Db 1859 ACCATTACTTCAATAGCAGCAAAAGCTCTTAGAAGAACATGTGCTTCATGATGATCCGG 1918
Qy 1809 AGTTCGGCGGAAGGTCTGCTACGTTCAAGTTCGCGCAGAGTTTCGACGGCATCAAGTGC 1868
Db 1919 CTCTAGGAAGGAAAACCTTGTTATGTATGTACAAATTTCCACAGAGATTTGATGGCAATTGACTTGC 1978
Qy 1869 ACAGCCGATACGCTAAACAGGAACACCGTCTTCTCGACATCAACATGAAGGGCTGGACG 1928
Db 1979 ACGATCGATATGCTAATCGGAACATAGTTTTCTTTTGATATCAACATGAAGGTCTGGATG 2038
Qy 1929 GCATCCAAAGGCCGCTGTACGTCCGGGACAGGGTGGTGTTCGGCGCCAGGCGCTCTACG 1988
Db 2039 GCATTCAAGGTTCCAGTTTACGTGGGAACAGGATGCTGTTTCAATAGACAGGCTTTGTATG 2098
Qy 1989 GCTACAAACCTT-----CCCAAGGGAACCAAGAGGCCAAGATGCTGACCTGCGACTGCT 2042
Db 2099 GATACGATCTCTGTTTGTGATGAAGTGAATCTGAGAGCCAAAATTTGTTATTAAGAGCTGCT 2158
Qy 2043 G-----CCCGTGTCTCGCGCGCAAGAGCGGAAGCGGAACACGCCAA 2079
Db 2159 GTGTGAGGAAGAAAGAAACAGAGTTATATGATAGTATCAAGCCGTTATTAAGA 2218
Qy 2080 GGACGGGTGC-----CGGAGGGCACCGCTTGATATGGAGTAGATA 2120
Db 2219 GAACAGAATCTTCAGTCCCATCTTCAATATGGAAGACATCGAAGAGGGTATTGAAGTT 2278
Qy 2121 GCGACAGGAGA-----TGCTCATGTCCACATGAATTCGAGAAGCGGTTTCGGGCAGT 2174
Db 2279 ACAGAGGATGAAGGTTCAGTGTCTTATGTCCCAAGAGGAAATTTGGAGAAACGCTTTGGTCAGT 2338
Qy 2175 CCGCGGCGTTGCTCAGCTCGACGCTGATGGAGGAGGGCGCTCCCTCTCTCGTGGAGCC 2234
Db 2339 CTCCTATTTCATTTGCATCCACCTTTTATGACAAAGGTGGCATACCACTTCAACAAAC 2398

Qy 2235 CCGCCCGCGCTCCTCAAGAGAGGCCATCCATGTATCAGTGGGCTACGAGACAAGACCG 2294
Db 2399 CAGCTTCTTACTAAAGGAAGCTATCCATGTCTCATCAGTTGTGGATATGAGGACAAAAC 2458
Qy 2295 ACTGGGGGCTGGAGCTCGGGGTGGATCTACGGGTGATCAGGAGGACATCTCTGACGGGT 2354
Db 2459 AATGGGAAAGAGATTGGCTGGATCTATGGTTCACTAACGAGAGATATCTGACTGGGT 2518
Qy 2355 TCAAGATGCACTGCGCGGGTGGCGCTCCGTGTACTGATGCCGAAGCGCGGGTTC 2414
Db 2519 TTAAGATGCATGCAAGGGCTGGCAATCAATCTACTGATGCCACCAAGCATCTTGTTC 2578
Qy 2415 AGGGGTGCGGCCCGATCAATCTATCGGACCGTCTCAACCGAGTGTCTCGGTGGGGCTGG 2474
Db 2579 AGGGTCTGCAACCAATCAATCTTTCCGATCGTCTTAATCAGGTGCTCCGTGGGGCTCTG 2638
Qy 2475 GGTCCGTCGAGATCTTCTCAGCGCGCACAGCCCCCTGCTGTACGGCTACAAGAACGCA 2534
Db 2639 GGTCAAGTGGAAATTTCTCTTAGTAGACATTTGCTCTATCTGTGTATGGTTTACAAT---GGAC 2695
Qy 2535 ACCTCAAGTGGCTGGAGCGCTTCGCCCTACATCAACACCAACCATCTACCCCTTCACCTCGC 2594
Db 2696 GATTGAAGCTTTTGGAGAGGCTGGCTTACATCAACACTATTGTATATCCATCACAATCCA 2755
Qy 2595 TCCCGCTGCTCGCTACTGCAACCTCCCGCGCTCTGCTCTCACGGGGAAGTTTCATCA 2654
Db 2756 TTCGCTTATTTGGCTATTGTGTGCTTCCCGCTATCTGCTCTTACCAATAAATTTATCA 2815
Qy 2655 TGCGTTCGATTAGCACGTTCCGACGCTTCTTCTCATCGCCCTCTTCTGTCATCTTCTCG 2714
Db 2816 TTCCTGAGATTAGCAAAATTTATGCTGGGATTTTCTTCATTTCTTCTTCTGCTTCTTCTG 2875
Qy 2715 CGAGCGGCATCTCTGGAGATCGGTGGAGCGGGGTGAGCATCGAGAGTGTGGAGGAACG 2774
Db 2876 CCACTGTATATTGGAGCTTAGATGGAGTGGTGTGGCATTTGAAGATTGGTGGAAATG 2935
Qy 2775 AGCAGTTCTGGGTCACTCGCGGGGTGCGGGCATCTCTTTCGCGCTGTGTGAGGGCTTGC 2834
Db 2936 AGCAGTTTGGGTATTGGTGGCACCTCTGCCCCATCTCTTCGCAAGTGTTCAGGGGTCTGC 2995
Qy 2935 TCAAGTCTCTCGCGGATCGACCAACATTTACCGTCACTCCCAAGGCCACCGCGCAGC 2894
Db 2996 TGAAGTGTGGCTGGGATTTGATACCAACTTCACAGTTTACCTCAAAGGCATC---TGATG 3052
Qy 2895 AGGACGACGAGTTCCGCGAGCTCTACGCCCTTCAAGTGGACCCACGCTCTCTCATCCGCCCA 2954
Db 3053 AGGATGGGACTTTGCTGAGCTATATGTTTCAAGTGGACCAAGTTTGTCTCATTTCTCGA 3112
Qy 2955 CCAGCTGTCTCATATTAAAGTCAATCGGCTGTGTGGCGGGATCTCCGACGCCATCAACA 3014
Db 3113 CCACTGTTCTTGTTCATTAACTGCTGGAATGGTGGCAGGAATTTCTTATGCCATTAACA 3172
Qy 3015 ACGGGTACCAAGTCTCTGGGGGCGCTTTCGCGAGCTTCTTCTGCGCTTCTGCGGTCTACG 3074
Db 3173 GTGGCTACCAATCTCTGGGGTCCGCTCTTTGGAAGCTGTCTTCTCTGATTTGGGTGATCC 3232
Qy 3075 TCCACCTCTACCCGCTTCTCAAGGGGCTCATGGGGCGCCAGAACAGACGCGCCACCGGTG 3134
Db 3233 TCCATCTTACCCCTTCTCAAGGGTCTCATGGAGGACGACGACACCAACATCG 3292
Qy 3135 TTGTCTATCGTTCATTTCTGCTGCTTCCATTTCTTCTCTGCTCTGGGTCAAGATTCGACC 3194
Db 3293 TCATTGTCTGCTTCCATCTTCTGCTATCTTCTCTGCTGCTGGGTGAAGATTCGATC 3352
Qy 3195 CTTTTCATCGTCAGAC---CAAGGGCCGAGCGTTCAGGCGATGAGGATCAATTTGCTGAG 3251
Db 3353 CTTTTCATCTCCCGACACAGAAAGCTGTGCTTGGGGCAATGTGGCGTCAACTGCTGAT 3412
Qy 3252 C 3252
Db 3413 C 3413

1749	Qy	ATCACTACATCAACAACAGCAAGGCCATTCGGGAGGCCATGTGCTTCTCATGGAACCTC	1808
1876	Db	ACCAATTACTTTCAATAGCAGCAAAAGCTCTTAGAGAAAGCAATGTGTTTCATCATGATGCATCGG	1935
1809	Qy	AGGTCCGCCCGGAAGGTCTGCTAGCTTCAGTTTCCCGCAGAGGTTTCGACGGCATCGACGTGC	1868
1936	Db	CTCTAGAGAAAGAAACTTTGTTATGTACAAATTTCCACAGAGATTTGATGGCAATTTGACTTGC	1995
1869	Qy	ACGACCGATACGCTAAACAGGAAACACCGTCTTCTTCGACATCAACATGAAGGGGCTGGACG	1928
1996	Db	ACGATCGATATGCTAATTCGGAACATAGTCTTTCTTTGATATCAACATGAAGGCTCGGATG	2055
1929	Qy	GCATCCAAAGCCCGGTGTAGTCTGGGACAGGTCGCTGTTCCGGCGCCAGCGGCTCTACG	1988
2056	Db	GCATTCAGGGTTCAGTTTACGTGGGAAACAGATGCTGTTTCAATAGACAGGCTTTGTATG	2115
1989	Qy	GCTACAAACCTT-----CCCAAGGAGCCCAAGAGGCCCAAGATGTTGACCTGCGACTGCT	2042
2116	Db	GATACGATCTGTTTGACTGAAAGTGAATCTGGAGCCAAACATTTGTTATTAAAGAGCTGCT	2175
2043	Qy	G-----CCCGTGTCTTCGGCCGCAAGAGCGGAAAAACACGCCAA	2079
2176	Db	GTGGTAGAAGGAAGAAAAGAAACAAGAGTTATATGGATAGTCAAAGCCGTATTATGAAGA	2235
2080	Qy	GGAACGGGCTGC-----CGAGGGGCAACGCTGTATATGGGAGTAGATA	2120
2236	Db	GAACAGAAATCTTCAGCTCCCATCTTCAATATGGAAGACATCGAAGAGGGTATTGAAGGTT	2295
2121	Qy	GCACACAGGAGA-----TGCTCATGTCCACATGACTTCGAGAGAGCGGTTTCGGGCAGT	2174
2296	Db	ACGAGGATGAAGGTCAGTGCTTATGTCCAGAGGAAATTTGGAGAAACGCTTTGGTCACT	2355
2175	Qy	CCGCGGGCTTCGTCACTCGACGCTGATCGAGGAAAGCGCGCGTCCCTCTTCCTCGTCGAGCC	2234
2356	Db	CTCTATTTCATTGCAATCCACTTTATGACACAAGGTGSCATACCACTTTCACAAACC	2415
2235	Qy	CCGCGCGCTCTCAAGGAGGCCATCCATGTCAATCAAGTCTCGGCTACGAGGACAAAGACG	2294
2416	Db	CAGCTTCTCTACTAAAGGAAGCTATCCATGTCACTAGTTGTGATATGAGAACAAACATG	2475
2295	Qy	ACTGGGGCTGGAGCTGGGTGGATCTACGGGTGATCAGGAGGACATCTTCACCGGGT	2354
2476	Db	AATGGGGAAGAGATTGGCTGCACTCATGGTTCAGTAACGGAAGGATTTCTCACTGGGT	2535
2355	Qy	TCAAGATGCACTGCGCGGTGGCGCTCCGTGTACTGCAATGCCAGACGGCGCGCTTCA	2414
2536	Db	TTAAATGATGATGAAAGGGCTGGCAATCAATCTACTGCAATGCCACCAAGACCTTGTTCA	2595
2415	Qy	AGGGTTCGGCGCGATCAATCTATCGGACCGTCTCAACAGGTGCTCCGTTGGCGCTGG	2474
2596	Db	AGGGTTCGCAACCAATCAATCTTTCCGATCGTCTTAATCAGGTGCTCCGTTGGGCTCTTG	2655
2475	Qy	GGTCCGTCGAGATCTTCTTCAGCGGACAGCCCCCTGCTGTACGGCTACAAAGACGGCA	2534
2656	Db	GGTCAGTGGAAATTTCTGCTTAGTAGACATTTGCTCTATCTGTTATGGTTACAAT---GGAC	2712
2535	Qy	ACCTCAAGTGGCTGGAGCGCTTGCCTPACATCAACACCAACCATCTACCCCTTCACCTCGC	2594
2713	Db	GATTTGAAGCTTTTGGAGAGGCTGGCTTACATCAACACTATTGTATATCCAAATCACATCCA	2772
2595	Qy	TCCCGCTGTCTCGCTACTGCACCCCTCCCGCGGTCTGCTCCTTCACCGGCAAGTTCATCA	2654
2773	Db	TTCCGCTTATTGCTATTGTGTGCTTCCCGCTATCTGCGCTCTTACCAATAAATTTATCA	2832
2655	Qy	TGCGGTGATTAGCAGTTTCGACAGCCTCTTCTTCATCGGCCCTTCTCATGTCCATCTTCG	2714
2833	Db	TTCTTGAGATTAGCAATTTATGCTGGGATGTCTTCAATCTTCTTTTCGCTCCATTTTTCG	2892
2715	Qy	CGACGGCATCTTCGGAGATCGGTGGAGCGGGTGGAGCATCGAGGAGTGGTGAGGAACG	2774
2893	Db	CCACTGGTATATTGGAGCTTAGATGGAGTGGTGTGGCAATTTGAAGATTGGTGAGAAATG	2952
2775	Qy	AGCAGTTCTGGGTCACTCGGCGGGGTGTTCGCGCATCTCTTCGCGCTGTCGACGGGCTGC	2834

Db	2953	AGCAGTTTGGGTATTGGTGGCACTCTGCGCACTCTTCGACAGTTTCCAGGGTCTCG	3012
Qy	2835	TCAAGTCTCGCGGGATCGACACCAACTTCAACGTCACCTCCAAAGGCACCGCGCAG	2894
Db	3013	TGAAAGTTGGCTGGGATTGATACCAACTTCACAGTTACCTCAAAGGCATC--TGATG	3069
Qy	2895	AGACAGCAGTTTCGGCGAGCTCTACGCCCTTCAAGTGGACGACGCTCTCATCCCGCCA	2954
Db	3070	AGGATGGCGACTTTTGCTGAGCTATATGTGTCAAGTGGACCAAGTTTGCTCATTCCTCCGA	3129
Qy	2955	CCAGCTGCTCATCATTTAAGTCATCGCGTGTGTCGCGGCATCTCCGACGCCATCAACA	3014
Db	3130	CCACTGTTCTTGTCATTAACTGGTCCGAATGGTGCAGAAATTTCTTAGCATTTAACA	3189
Qy	3015	ACGGGTACCAAGTCTTCGGGGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTCAATCG	3074
Db	3190	GTGGCTACCAACTCCTGGGGTCGCTCTTTGGAAAGCTGTTCTTCGATCTGGGTATCC	3249
Qy	3075	TCCACTCTACCGTTCTCTCAAGGGGCTCATGGGGCGCAGAACAGGACGCCACCGTTG	3134
Db	3250	TCCATCTCTACCCCTTCTCAAGGGTCTCATGGGAAGGCAGAACCGCACACCAACAATCG	3309
Qy	3135	TTGTCATCTGGTCCATTCTGCTGGCCCTCCATCTTCTCCCTGCTCTGGGTCAAGTCGACC	3194
Db	3310	TCATTGTCGGTCCATCTCTTTCGATCTATCTTCTCTGCTGGGTGAAGATCGATC	3369
Qy	3195	CTTTTCATCGTCAAGAC---CAAGGGGCCGAGCGTCAGGCAGTGTGGCATCAATTGCTGAG	3251
Db	3370	CTTTTCATCTCCCGACACAGAAAGCTGCTGCCCTTGGGGCAATGTGGCGTCAACTGCTGAT	3429
Qy	3252	C 3252	
Db	3430	C 3430	

Search completed: March 12, 2005, 00:02:03
Job time : 425 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 21:58:45 ; Search time 1220 Seconds
(without alignments)
16787.522 Million cell updates/sec

Title: US-10-627-132-29

Perfect score: 3443
Sequence: 1 ctgctgcctgcctgcga.....aaaaaaaaaaaaaaaa 3443

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3443	100.0	3443	17	US-10-627-132-29
2	3028	87.9	3028	16	US-10-209-059-29
3	3008.6	87.4	3589	17	US-10-425-114-24930
4	2692.8	78.2	3448	18	US-10-437-963-15093
5	2633.2	76.5	2687	17	US-10-425-114-833
6	2495.8	72.5	2830	18	US-10-425-115-139789
7	1563.4	45.4	3470	16	US-10-209-059-25
8	1563.4	45.4	3470	17	US-10-627-132-25
9	1523.2	44.2	3517	9	US-09-900-237-13
10	1439.6	41.8	3493	18	US-10-437-963-83811
11	1397.6	40.6	3746	16	US-10-160-719-29

12	1397.6	40.6	3746	16	US-10-160-719-49	Sequence 49, Appl
13	1397.6	40.6	3773	16	US-10-209-059-9	Sequence 9, Appl
14	1397.6	40.6	3773	16	US-10-160-719-5	Sequence 5, Appl
15	1397.6	40.6	3773	17	US-10-627-132-9	Sequence 9, Appl
16	1382.4	40.2	3854	18	US-10-425-115-102948	Sequence 102948,
17	1368.6	39.8	3851	17	US-10-393-840-69	Sequence 69, Appl
18	1366.4	39.7	3987	18	US-10-425-115-45192	Sequence 45192, A
19	1364.8	39.6	3786	9	US-09-900-237-7	Sequence 7, Appl
20	1364.8	39.6	3813	16	US-10-209-059-17	Sequence 17, Appl
21	1364.8	39.6	3813	16	US-10-160-719-25	Sequence 25, Appl
22	1364.8	39.6	3813	16	US-10-160-719-45	Sequence 45, Appl
23	1364.8	39.6	3813	17	US-10-627-132-17	Sequence 17, Appl
24	1363.2	39.6	3814	18	US-10-425-115-45185	Sequence 45185, A
25	1325	38.5	3812	18	US-10-425-115-101245	Sequence 101245,
26	1323.4	38.4	3799	16	US-10-267-459-5	Sequence 5, Appl
27	1323.4	38.4	3799	16	US-10-209-059-21	Sequence 21, Appl
28	1323.4	38.4	3799	17	US-10-627-132-21	Sequence 21, Appl
29	1315	38.2	2684	18	US-10-425-115-137279	Sequence 137279
30	1268	36.8	3776	9	US-09-900-237-3	Sequence 3, Appl
31	1266.4	36.8	3505	17	US-10-425-114-31130	Sequence 31130, A
32	1266.4	36.8	3783	18	US-10-425-115-86998	Sequence 86998, A
33	1245.2	36.2	3626	9	US-09-900-237-29	Sequence 29, Appl
34	1242.4	36.1	3962	18	US-10-437-963-72402	Sequence 72402, A
35	1233.2	35.8	3286	18	US-10-425-115-43488	Sequence 43488, A
36	1230.2	35.7	2758	17	US-10-424-599-125586	Sequence 125586,
37	1222.4	35.5	3568	16	US-10-209-059-41	Sequence 41, Appl
38	1222.4	35.5	3568	16	US-10-160-719-1	Sequence 1, Appl
39	1222.4	35.5	3568	17	US-10-627-132-41	Sequence 41, Appl
40	1211.6	35.2	3744	18	US-10-425-115-100413	Sequence 100413,
41	1208.4	35.1	3704	16	US-10-209-059-13	Sequence 13, Appl
42	1208.4	35.1	3704	16	US-10-160-719-57	Sequence 57, Appl
43	1208.4	35.1	3704	17	US-10-627-132-13	Sequence 13, Appl
44	1204.6	35.0	3936	9	US-09-900-237-9	Sequence 9, Appl
45	1204.6	35.0	3969	16	US-10-209-059-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1

US-10-627-132-29
; Sequence 29, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-29

Query Match 100.0%; Score 3443; DB 17; Length 3443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCGTGCCTGCCTGCATCGGAATCTGTGAGCACCTGAGGGTCCGAGCCGAG 60
|||||

Db 1 CTGCGTCGCCCTCGCTCGCAATCCGGAACTCTGTGAGACACCTGAGGGTCCGAGGCCGAG 60
Qy 61 AGCTAGCTAGCACGCCGCGCTCCGGCGCGAGATGAGAGCCAGCGCCGGGCTGCTGGCCGG 120
Db 61 AGCTAGCTAGCACGCCGCGCTCCGGCGCGAGATGAGAGCCAGCGCCGGGCTGCTGGCCGG 120
Qy 121 CTGCGCAACCCGGAACGAGTGTGTGTGTATCCGGGGCCACGAGGACCCCAAGCCGTGGG 180
Db 121 CTGCGCAACCCGGAACGAGTGTGTGTGTATCCGGGGCCACGAGGACCCCAAGCCGTGGG 180
Qy 181 GGGCTGAGCGGCGAGGTGCGAGATATGCGCGACGAGGTGCGGCTCACGCTGAGCCG 240
Db 181 GGGCTGAGCGGCGAGGTGCGAGATATGCGCGACGAGGTGCGGCTCACGCTGAGCCG 240
Qy 241 CGACCTCTTCGTCGCTCGCAACGAGTCCGCTTCCCGTGTGCGGCCCTGCTACGAGTA 300
Db 241 CGACCTCTTCGTCGCTCGCAACGAGTCCGCTTCCCGTGTGCGGCCCTGCTACGAGTA 300
Qy 301 CGAGCGCGGGAGGGGACGAGAACTGCCCCAGTGCAGAGACGCGCTACAAGCGCTCAA 360
Db 301 CGAGCGCGGGAGGGGACGAGAACTGCCCCAGTGCAGAGACGCGCTACAAGCGCTCAA 360
Qy 361 GGGGAGCCCGAGGGTTGCCGGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGA 420
Db 361 GGGGAGCCCGAGGGTTGCCGGGACGATGACGAGGAGGAGGACATCGACGACCTGGAGCACGA 420
Qy 421 GTTCAAATCGACGACGAGAAATCAGCAGAGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGA 480
Db 421 GTTCAAATCGACGACGAGAAATCAGCAGAGGACGATGAGGAGGACATCGACGAGGACGACCA 480
Qy 481 GATCACCGAGCGATGTGCA CGGCAAGATGAGCTACGGAGGGGCGCGACGAGCGGA 540
Db 481 GATCACCGAGCGATGTGCA CGGCAAGATGAGCTACGGAGGGGCGCGCGACGAGCGGA 540
Qy 541 CGGCAACACACCCCGAGATCCCGCCCATCATCACCGCTCCCGTCCGCTGCGCGTGAG 600
Db 541 CGGCAACACACCCCGAGATCCCGCCCATCATCACCGCTCCCGTCCGCTGCGCGTGAG 600
Qy 601 CGGTGAGTTTCCGATTTACCAACGGGTATGGCCACGGCGAGGTCTCGTCTTTCCCTGCACAA 660
Db 601 CGGTGAGTTTCCGATTTACCAACGGGTATGGCCACGGCGAGGTCTCGTCTTTCCCTGCACAA 660
Qy 661 GGGCATCCATCCGTACCTGTGTCTGAGCCAGGAGTGCCAAAGTGGGA GAGAAAGA 720
Db 661 GGGCATCCATCCGTACCTGTGTCTGAGCCAGGAGTGCCAAAGTGGGA GAGAAAGA 720
Qy 721 AGTGAGCTGGAAGGAGGATGAGCACTGGAAGTCCAAAGCAGGGCATCTCTGGCGCGG 780
Db 721 AGTGAGCTGGAAGGAGGATGAGCACTGGAAGTCCAAAGCAGGGCATCTCTGGCGCGG 780
Qy 781 CGCGATCCCGAAGACATGAGACCCGACGCTGGCACTGAAACGAGCGAGGCGAGCCGCT 840
Db 781 CGCGATCCCGAAGACATGAGACCCGACGCTGGCACTGAAACGAGCGAGGCGAGCCGCT 840
Qy 841 GTCGAGGAAGGTGTCATCCGCTCGAGCAAGGTGAACCCGTA CCGAATGGTGAATCGTGGT 900
Db 841 GTCGAGGAAGGTGTCATCCGCTCGAGCAAGGTGAACCCGTA CCGAATGGTGAATCGTGGT 900
Qy 901 GGGTCTGCTGTGCTCGCTTCTTCTCCGTTACGTTATCTGCA CCCCCTCCCGACCG 960
Db 901 GGGTCTGCTGTGCTCGCTTCTTCTCCGTTACGTTATCTGCA CCCCCTCCCGACCG 960
Qy 961 CATCGGGCTGTGGCTCGTCTCATCATCTGCGAGATCTGGTTCGCCATCTCTGGATCCT 1020
Db 961 CATCGGGCTGTGGCTCGTCTCATCATCTGCGAGATCTGGTTCGCCATCTCTGGATCCT 1020
Qy 1021 CGACCAAGTCCCAAGTGGTTCCCAATCGACCGGAGACGTACTCGACCGCTCTCCCT 1080
Db 1021 CGACCAAGTCCCAAGTGGTTCCCAATCGACCGGAGACGTACTCGACCGCTCTCCCT 1080
Qy 1081 CAGGTACGAGAGGAAGGGAGCCGCTGCTGCTGTCGCGGCTGAGACCTGTTGTGAGCAC 1140
Db 1081 CAGGTACGAGAGGAAGGGAGCCGCTGCTGCTGTCGCGGCTGAGACCTGTTGTGAGCAC 1140

Qy 1141 GGTGGAACCGCTCAAGGAGCGCGCTGTGTGACCGCAACACCGCTGTCTTCCATCTCTCGC 1200
Db 1141 GGTGGAACCGCTCAAGGAGCGCGCTGTGTGACCGCAACACCGCTGTCTTCCATCTCTCGC 1200
Qy 1201 CGTAGACTACCCCGTGGACAAAGGTCTCTGCTAGTCTCCGACGAGCGCGCTGCGATGCT 1260
Db 1201 CGTAGACTACCCCGTGGACAAAGGTCTCTGCTAGTCTCCGACGAGCGCGCTGCGATGCT 1260
Qy 1261 GAGCTTCGAGTCTGCTCGGAGACGCGCGAGTTCGCGCGCAAGTGGTGCTTCTGCAA 1320
Db 1261 GAGCTTCGAGTCTGCTCGGAGACGCGCGAGTTCGCGCGCAAGTGGTGCTTCTGCAA 1320
Qy 1321 GAAGTTCGGCATCGAGCCCCCGCGCGCGAGTTCTACTTCTCGCTCAAGGTGCGACTACCT 1380
Db 1321 GAAGTTCGGCATCGAGCCCCCGCGCGCGAGTTCTACTTCTCGCTCAAGGTGCGACTACCT 1380
Qy 1381 CAAGGACAAGGTGACGCCCACTTCTGTGACGAGCGCGCGCATGAAGAGAGATGA 1440
Db 1381 CAAGGACAAGGTGACGCCCACTTCTGTGACGAGCGCGCGCATGAAGAGAGATGA 1440
Qy 1441 GGAGTTCAAGGTCCGGATCAA CGCGCTGTGTGGCCAAAGGCCATGAAGTGCAGAGGG 1500
Db 1441 GGAGTTCAAGGTCCGGATCAA CGCGCTGTGTGGCCAAAGGCCATGAAGTGCAGAGGG 1500
Qy 1501 GTGGATCATGAAGGACGCGCACCGCTGGCCCGGGAAACAACACCCCGGCGCAT 1560
Db 1501 GTGGATCATGAAGGACGCGCACCGCTGGCCCGGGAAACAACACCCCGGCGCAT 1560
Qy 1561 GATCAGAGTGTCTCGGCCACAGCGCGGCCACGACGAGGGCAACGAGCTGCCCG 1620
Db 1561 GATCAGAGTGTCTCGGCCACAGCGCGGCCACGACGAGGGCAACGAGCTGCCCG 1620
Qy 1621 CCTCGTACGCTCTCCGTGAGAGCGCGCGGATTCACAGCACCAAGAGGCCGCGC 1680
Db 1621 CCTCGTACGCTCTCCGTGAGAGCGCGCGGATTCACAGCACCAAGAGGCCGCGC 1680
Qy 1681 CATGAACGCTCTGATTCGGCTCTCCCGCTGTGTGACCAACCGCCCATTCATGTCTCAACTT 1740
Db 1681 CATGAACGCTCTGATTCGGCTCTCCCGCTGTGTGACCAACCGCCCATTCATGTCTCAACTT 1740
Qy 1741 GGACTGTGATCACTACATCAACAAAGGCGCATCCGGGAGGCGCATGTGCTTCTCAT 1800
Db 1741 GGACTGTGATCACTACATCAACAAAGGCGCATCCGGGAGGCGCATGTGCTTCTCAT 1800
Qy 1801 GGAACCTCAGGTCCGGCGGAGGTCTGCTAGTTCAGTTCCTCCGACAGGTTCCAGCGCAT 1860
Db 1801 GGAACCTCAGGTCCGGCGGAGGTCTGCTAGTTCAGTTCCTCCGACAGGTTCCAGCGCAT 1860
Qy 1861 CGACGTGCA CGACCGATACGCTAAACAGGAACACCGTCTTCTTCGACATCAACATGAAGG 1920
Db 1861 CGACGTGCA CGACCGATACGCTAAACAGGAACACCGTCTTCTTCGACATCAACATGAAGG 1920
Qy 1921 GCTGAGCGCATCCAAAGGCCCGGTGTACGTCCGGGACAGGGTGTGTTCCGGCGCCAGGC 1980
Db 1921 GCTGAGCGCATCCAAAGGCCCGGTGTACGTCCGGGACAGGGTGTGTTCCGGCGCCAGGC 1980
Qy 1981 GCTCTACGCTACAAACCTCCCAAGGACCCCAAGAGGCCCAAGATGATGCTGCGACTG 2040
Db 1981 GCTCTACGCTACAAACCTCCCAAGGACCCCAAGAGGCCCAAGATGATGCTGCGACTG 2040
Qy 2041 CTGCCGCTCTTCGGCGCAAGAGCGGAACACGCAAGGAGCGGCTGCGGAGGGCAC 2100
Db 2041 CTGCCGCTCTTCGGCGCAAGAGCGGAACACGCAAGGAGCGGCTGCGGAGGGCAC 2100
Qy 2101 CGCTGATATGGAGATGATAGCGA CAAAGAGATGCTCATGTCCCAATGAACCTTCGAGAA 2160
Db 2101 CGCTGATATGGAGATGATAGCGA CAAAGAGATGCTCATGTCCCAATGAACCTTCGAGAA 2160
Qy 2161 GGGGTTCCGGGACGTCGCGCGGCTTCGTACGCTCGAGGAGGAGGCGCGCTCCC 2220
Db 2161 GGGGTTCCGGGACGTCGCGCGGCTTCGTACGCTCGAGGAGGAGGCGCGCTCCC 2220


```
Qy 2221 TCCTTCGTGAGCCCGCGCGCTCCTCAAGAGGCCATCATGTCTATGCTGAGCTCGCGGCTA 2280
Db 2221 TCCTTCGTGAGCCCGCGCGCTCCTCAAGAGGCCATCATGTCTATGCTGAGCTCGCGGCTA 2280
Qy 2281 CGAGCAAGACCGACTGGGGCTGGAGCTGGGTGGATCTACGGGTGATCAACGAGGA 2340
Db 2281 CGAGCAAGACCGACTGGGGCTGGAGCTGGGTGGATCTACGGGTGATCAACGAGGA 2340
Qy 2341 CATCTGAGCGGGTTCAAGATGCACTGCGCGGGTGGGTCTCGTGTACTGCAATGCCGAA 2400
Db 2341 CATCTGAGCGGGTTCAAGATGCACTGCGCGGGTGGGTCTCGTGTACTGCAATGCCGAA 2400
Qy 2401 GCGGCGCGCTTCAAGGGTCCGCGCGATCAATCTATCGGCGCTCTCAACGAGTCT 2460
Db 2401 GCGGCGCGCTTCAAGGGTCCGCGCGATCAATCTATCGGCGCTCTCAACGAGTCT 2460
Qy 2461 CCGGTGGCGCTGGGGTCCGTCGAGATCTTCTTCAGCGCGCACAGCCCTCTGTACGG 2520
Db 2461 CCGGTGGCGCTGGGGTCCGTCGAGATCTTCTTCAGCGCGCACAGCCCTCTGTACGG 2520
Qy 2521 CTACAGAAAGGCAACTCAAGTGGCTGGAGCGCTTCGCTTACATCAACACCACTCTA 2580
Db 2521 CTACAGAAAGGCAACTCAAGTGGCTGGAGCGCTTCGCTTACATCAACACCACTCTA 2580
Qy 2581 CCCCTTCACTCGCTCGCTGCTGCTGCTACTGCACTCCCGCGCTGCTGCTCTCAC 2640
Db 2581 CCCCTTCACTCGCTCGCTGCTGCTGCTACTGCACTCCCGCGCTGCTGCTCTCAC 2640
Qy 2641 CGGCAAGTTCATCATGCGCTGATAGCATGTTCCGACGCTTCCTTCATCGCCCTT 2700
Db 2641 CGGCAAGTTCATCATGCGCTGATAGCATGTTCCGACGCTTCCTTCATCGCCCTT 2700
Qy 2701 CATGTCCATCTTCGCGAGGCACTCTGGAGATCGGTGGAGCGGGTGAGCATCGAGGA 2760
Db 2701 CATGTCCATCTTCGCGAGGCACTCTGGAGATCGGTGGAGCGGGTGAGCATCGAGGA 2760
Qy 2761 GTGGTGAAGAACGAGCAGTCTTGGGTCTATCGGCGGGTGTCCGCGCATCTCTTCGCGT 2820
Db 2761 GTGGTGAAGAACGAGCAGTCTTGGGTCTATCGGCGGGTGTCCGCGCATCTCTTCGCGT 2820
Qy 2821 CGTGAGGGCTGTCTCAAGTCTCTCGCGGGATCGACCAACTTCAACGCTCACTCCAA 2880
Db 2821 CGTGAGGGCTGTCTCAAGTCTCTCGCGGGATCGACCAACTTCAACGCTCACTCCAA 2880
Qy 2881 GGCACCGGCGACGAGACGAGTTCGCGAGCTCTACGCTTCAAGTGGACCACTCT 2940
Db 2881 GGCACCGGCGACGAGACGAGTTCGCGAGCTCTACGCTTCAAGTGGACCACTCT 2940
Qy 2941 CCTCATCCCGCCACACGCTGCTCATCATTAAGTCTATCGGCGTCTGCGCGCATCTC 3000
Db 2941 CCTCATCCCGCCACACGCTGCTCATCATTAAGTCTATCGGCGTCTGCGCGCATCTC 3000
Qy 3001 CGAGCCATCAACAAACGGGTACAGTCTCTGGGGCCCTCTTGGCAAGCTCTTCTTCG 3060
Db 3001 CGAGCCATCAACAAACGGGTACAGTCTCTGGGGCCCTCTTGGCAAGCTCTTCTTCG 3060
Qy 3061 CTTCTGGGTCTATCGTCCACTCTACCGGTTCTCAAGGGGCTCATGGGGCGCCAGAAC 3120
Db 3061 CTTCTGGGTCTATCGTCCACTCTACCGGTTCTCAAGGGGCTCATGGGGCGCCAGAAC 3120
Qy 3121 GAGCCACCGTGTGTCATCTGGTCCATCTGCTGGCCCTCAATCTTCTCCCTGCTG 3180
Db 3121 GAGCCACCGTGTGTCATCTGGTCCATCTGCTGGCCCTCAATCTTCTCCCTGCTG 3180
Qy 3181 GGTGAGATCGACCTTTTCAAGGTCAGGACCAAGGGCCGAGCTCAGGCGAGTGTGCGAT 3240
Db 3181 GGTGAGATCGACCTTTTCAAGGTCAGGACCAAGGGCCGAGCTCAGGCGAGTGTGCGAT 3240
Qy 3241 CAATTGCTGAGCTGTTTATTAAGGTTCAAAATTTCTGGAGCTTGTGATAGGAGAAAA 3300
Db 3241 CAATTGCTGAGCTGTTTATTAAGGTTCAAAATTTCTGGAGCTTGTGATAGGAGAAAA 3300
Qy 3301 ACAATTTAGAAATTTTGTAGGTTGTTGTCTGTGTAATGTTATGGTACCAGAAATGTCG 3360
```

```
Db 3301 ACAATTTAGAAATTTTGTAGGTTGTTGTCTGTGTAATGTTATGGTACCAGAAATGTCG 3360
Qy 3361 GACGAGCAATTTGAACAAAGGACAAAGTTTGTATTTGTTAAATGGCAAAAAA 3420
Db 3361 GACGAGCAATTTGAACAAAGGACAAAGTTTGTATTTGTTAAATGGCAAAAAA 3420
Qy 3421 AAAAAA 3443
Db 3421 AAAAAA 3443

RESULT 2
US-10-209-059-29
; Sequence 29, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3028
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-29

Query Match 87.9%; Score 3028; DB 16; Length 3028;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3028; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 416 CACGAGTTCAACATCGACGACGAGAAATCAGCAGGAGCTGGAGGGCAATCGCAGAAC 475
Db 1 CACGAGTTCAACATCGACGACGAGAAATCAGCAGGAGCTGGAGGGCAATCGCAGAAC 60
Qy 476 AGCCAGATCACGAGGAGTGTGCGAGGAGTGTGCTACGAGGAGGCGGCGGAGCAG 535
Db 61 AGCCAGATCACGAGGAGTGTGCGAGGAGTGTGCTACGAGGAGGCGGCGGAGCAG 120
Qy 536 GCGCAGCGCAACAAACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGTCCGTCGCG 595
Db 121 GCGCAGCGCAACAAACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGTCCGTCGCG 180
Qy 596 GTGAGCGGTGAGTTTCCGATTACCAACGGGTATGGCCACGCGGAGGTCTGCTTCCCTG 655
Db 181 GTGAGCGGTGAGTTTCCGATTACCAACGGGTATGGCCACGCGGAGGTCTGCTTCCCTG 240
Qy 656 CACAAGCGCATCCATCGTACCTGCTGTCTGAGCCAGGAGTGCAGTGGGACGAGAG 715
Db 241 CACAAGCGCATCCATCGTACCTGCTGTCTGAGCCAGGAGTGCAGTGGGACGAGAG 300
Qy 716 AAAGAGTGTAGCTGGAAGGAGAGATGGACGACTGGAAGTCCAAGCAGGAGCATCTCGGC 775
Db 301 AAAGAGTGTAGCTGGAAGGAGAGATGGACGACTGGAAGTCCAAGCAGGAGCATCTCGGC 360
Qy 776 GCGGCGCGCATCCGAAAGCATGAGCAGCGCATGCTGGCACTGAAACGACGAGGAGGAG 835
Db 361 GCGGCGCGCATCCGAAAGCATGAGCAGCGCATGCTGGCACTGAAACGACGAGGAGGAG 420
Qy 836 CCGCTGCGAGGAGGAGTGTGATCGGTCGAGCAAGTGAACCGTACCCGATGCTGATC 895
Db 421 CCGCTGCGAGGAGGAGTGTGATCGGTCGAGCAAGTGAACCGTACCCGATGCTGATC 480
```

QY	896	GTGGTGGCTCTCGTTGTGCTCGCCTTCTTCTCTCCGTAACGTAATCTGTACACCCCGTCCG	955
Db	481	GTGGTGGCTCTCGTTGTGCTCGCCTTCTTCTCTCCGTAACGTAATCTGTACACCCCGTCCG	540
QY	956	GAGCCATCGGCTGTGGCTGTCTCCATCATCTGCGAGATCTGGTTCGCCATCTCTCTG	1015
Db	541	GAGCCATCGGCTGTGGCTGTCTCCATCATCTGCGAGATCTGGTTCGCCATCTCTCTG	600
QY	1016	ATCCTCGACAGTTCCCAAGTGGTTCCCATCGACCGGAGAGTACCTCGACCGCCTC	1075
Db	601	ATCCTCGACAGTTCCCAAGTGGTTCCCATCGACCGGAGAGTACCTCGACCGCCTC	660
QY	1076	TCCCTCAGTACGAGAGGAGGGAGCCGTCTCTCTGTGCGCGGTGAGACCTGTTCGTG	1135
Db	661	TCCCTCAGTACGAGAGGAGGGAGCCGTCTCTCTGTGCGCGGTGAGACCTGTTCGTG	720
QY	1136	AGCACGGTGGACCGCTCAAGGAGCCGCGCTGGTGACCGCCACACCGTCTCTCATC	1195
Db	721	AGCACGGTGGACCGCTCAAGGAGCCGCGCTGGTGACCGCCACACCGTCTCTCATC	780
QY	1196	CTCGCGTGTAGACTACCCGTGGACAAGTCTCTCTGTCTCTCGACGACGCGCGTGC	1255
Db	781	CTCGCGTGTAGACTACCCGTGGACAAGTCTCTCTGTCTCTCGACGACGCGCGTGC	840
QY	1256	ATGCTGACGTTCTGAGTCTGTCGAGACGCGCGAGTTTCGCGCGAAGTGGTGCCTTC	1315
Db	841	ATGCTGACGTTCTGAGTCTGTCGAGACGCGCGAGTTTCGCGCGAAGTGGTGCCTTC	900
QY	1316	TGCAAGAAGTTCCGATCGAGCCCGCGCGCGCGAGTTCTACTTCTCGTCAAGTTCGAC	1375
Db	901	TGCAAGAAGTTCCGATCGAGCCCGCGCGCGCGAGTTCTACTTCTCGTCAAGTTCGAC	960
QY	1376	TACTCAAGGACAAGGTGCAAGCCACCTTCGTGCAAGAGCGCGCGCATGAAGAGAG	1435
Db	961	TACTCAAGGACAAGGTGCAAGCCACCTTCGTGCAAGAGCGCGCGCATGAAGAGAG	1020
QY	1436	TATGAGGAGTCAAGGTCCGATCAACGCGCTGGTGGCCAAAGGCCATGAAGTGC CGGCA	1495
Db	1021	TATGAGGAGTCAAGGTCCGATCAACGCGCTGGTGGCCAAAGGCCATGAAGTGC CGGCA	1080
QY	1496	GAGGGTGGATCATGAAGACGCGACCGCGTGGCGCGGAGCAACACCGCGACACCCC	1555
Db	1081	GAGGGTGGATCATGAAGACGCGACCGCGTGGCGCGGAGCAACACCGCGACACCCC	1140
QY	1556	GGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCGCACACACCGAGGCAACGAGCTG	1615
Db	1141	GGCATGATCCAGGTGTTCTTGGGCCACAGCGCGGCGCACACACCGAGGCAACGAGCTG	1200
QY	1616	CCCCGCTCGTGTACGTTCCCTGAGAGCGCCCGGGATTCAGCAACAAGAAAGGCC	1675
Db	1201	CCCCGCTCGTGTACGTTCCCTGAGAGCGCCCGGGATTCAGCAACAAGAAAGGCC	1260
QY	1676	GGCGCATGACCGCTCTGATTCGGTCTCGCGCTGTGACCAACCGGCCATTCATGCTC	1735
Db	1261	GGCGCATGACCGCTCTGATTCGGTCTCGCGCTGTGACCAACCGGCCATTCATGCTC	1320
QY	1736	AACTTGACATGTCATCACTACATCAACAGCAAGGCCATTCGGGAGGCCATGTGCTTC	1795
Db	1321	AACTTGACATGTCATCACTACATCAACAGCAAGGCCATTCGGGAGGCCATGTGCTTC	1380
QY	1796	CTCATGGACCTCAGGTTCGCGCGGAAGGTCTGCTACGTTCCCGCAGAGGTTTCGAC	1855
Db	1381	CTCATGGACCTCAGGTTCGCGCGGAAGGTCTGCTACGTTCCCGCAGAGGTTTCGAC	1440
QY	1856	GGCATCGACGTGACGACCGATAGCTAAACAGAAACACCGTCTTCTTCGACATCAACATG	1915
Db	1441	GGCATCGACGTGACGACCGATAGCTAAACAGAAACACCGTCTTCTTCGACATCAACATG	1500
QY	1916	AAGGGGCTGAGCGCATCCAAAGCCCGGTGTACGTTCGGGACAGGGTTCGTTCCCGCGC	1975
Db	1501	AAGGGGCTGAGCGCATCCAAAGCCCGGTGTACGTTCGGGACAGGGTTCGTTCCCGCGC	1560
QY	1976	CHAGGCGCTCTACGGCTACAAACCTCCCAAGGAGCCCAAGATGCTGACCTGC	2035
Db	1561	CHAGGCGCTCTACGGCTACAAACCTCCCAAGGAGCCCAAGATGCTGACCTGC	1620
QY	2036	GACTCTGCGCGTCTTCGGCGCGCAAGAACCGCAAGACGCGGCTGCGCGAG	2095
Db	1621	GACTCTGCGCGTCTTCGGCGCGCAAGAACCGCAAGACGCGGCTGCGCGAG	1680
QY	2096	GGCACCGCTGATATGGGAGTAGTAGCGACAAGAGAGTCTCATGTGCCACATGAATTC	2155
Db	1681	GGCACCGCTGATATGGGAGTAGTAGCGACAAGAGAGTCTCATGTGCCACATGAATTC	1740
QY	2156	GAGAAAGCGTTCCGGAGTTCGCGCGTTCGTCACTGTCAGCTGATGAGAGAGCGCG	2215
Db	1741	GAGAAAGCGTTCCGGAGTTCGCGCGTTCGTCACTGTCAGCTGATGAGAGAGCGCG	1800
QY	2216	GTCCCTCTCTTCGTCGAGCGCCCGCGCTCCTCAAGGAGCCATCATGTCTCATCTGC	2275
Db	1801	GTCCCTCTCTTCGTCGAGCGCCCGCGCTCCTCAAGGAGCCATCATGTCTCATCTGC	1860
QY	2276	GGCTACGAGGACAAGACCGACTGGGGGCTGAGCTGGGGTGGATCTACGGGTCTGATCAG	2335
Db	1861	GGCTACGAGGACAAGACCGACTGGGGGCTGAGCTGGGGTGGATCTACGGGTCTGATCAG	1920
QY	2336	GAGGACATCTGACGGGGTTCAGATGACCTCGCGCGGTGGCGTTCGTGTACTGCTCATG	2395
Db	1921	GAGGACATCTGACGGGGTTCAGATGACCTCGCGCGGTGGCGTTCGTGTACTGCTCATG	1980
QY	2396	CCGAAAGCGGCGGCTTCAAGGGTTCGCGCGGCTGAGCTGGGGTGGATCTACAGCAG	2455
Db	1981	CCGAAAGCGGCGGCTTCAAGGGTTCGCGCGGCTGAGCTGGGGTGGATCTACAGCAG	2040
QY	2456	GTGCTCCGGTGGCGGTGGGGTCCGTGAGATCTTCTTCAGCGGCGACACCCCTCTGCTG	2515
Db	2041	GTGCTCCGGTGGCGGTGGGGTCCGTGAGATCTTCTTCAGCGGCGACACCCCTCTGCTG	2100
QY	2516	TAAGGCTACAGAAACCGCAACCTCAAGTGGCTGAGAGCTTCGCTACATCAACACACC	2575
Db	2101	TAAGGCTACAGAAACCGCAACCTCAAGTGGCTGAGAGCTTCGCTACATCAACACACC	2160
QY	2576	ATCTACCCCTTCACTCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2635
Db	2161	ATCTACCCCTTCACTCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2220
QY	2636	CTCACGGCAAGTTTCATCATGCTCGATTAGACGCTTCGCGAGCTTCCTTTCATCGCC	2695
Db	2221	CTCACGGCAAGTTTCATCATGCTCGATTAGACGCTTCGCGAGCTTCCTTTCATCGCC	2280
QY	2696	CTTTCATGTCATCTTCGCGAGCGGCATCTGAGATCGGGTGGAGCGGGTGGAGCATC	2755
Db	2281	CTTTCATGTCATCTTCGCGAGCGGCATCTGAGATCGGGTGGAGCGGGTGGAGCATC	2340
QY	2756	GAGGAGTGTGGAGGAACGAGCAGTTCTGGGTTCATCGCGGGTTCGCGCATCTCTTC	2815
Db	2341	GAGGAGTGTGGAGGAACGAGCAGTTCTGGGTTCATCGCGGGTTCGCGCATCTCTTC	2400
QY	2816	GCGCTCGTCAGGGCTGCTCAAGGTCTTCGCGGGATCGACACCAACTTCACCGTCACC	2875
Db	2401	GCGCTCGTCAGGGCTGCTCAAGGTCTTCGCGGGATCGACACCAACTTCACCGTCACC	2460
QY	2876	TCCAAAGGCCACCGCGGACGAGGATTCGCGGCTTCACGCTTCACAGTGGAGC	2935
Db	2461	TCCAAAGGCCACCGCGGACGAGGATTCGCGGCTTCACGCTTCACAGTGGAGC	2520
QY	2936	ACGCTCTCATCTCCGCCACCGCTTCATCATTAAGCTTCATCGGCTTCGTTGGCGGGC	2995
Db	2521	ACGCTCTCATCTCCGCCACCGCTTCATCATTAAGCTTCATCGGCTTCGTTGGCGGGC	2580
QY	2996	ATCTCCGACGCCATCAACAGCGGTACAGTCTCGGGGCGCTCTTCGCGCAAGCTCTTC	3055
Db	2581	ATCTCCGACGCCATCAACAGCGGTACAGTCTCGGGGCGCTCTTCGCGCAAGCTCTTC	2640
QY	3056	TTCCGCTTCTGGGTTCATCTACCTCTACCGGCTTCCTCAAGGGGCTCATGGGGCGCCAG	3115

```
Db 2641 TTTGGCTTCTGGTTCATCGTCAACCTCTACCGTTTCTCAAGGGGCTCATGGGGCGCCAG 2700
Qy 3116 AACAGGACCCACCGTTGTTGTCACTCTGGTCCATTTCTGTGGCTCCATCTTCTCCCTG 3175
Db 2701 AACAGGACCCACCGTTGTTGTCACTCTGGTCCATTTCTGTGGCTCCATCTTCTCCCTG 2760
Qy 3176 CTCTGGGTTCAGATCGACCCCTTTCATCGTCAGGACCAAGGGCCGGAGCTCAGGCAGTGT 3235
Db 2761 CTCTGGGTTCAGATCGACCCCTTTCATCGTCAGGACCAAGGGCCGGAGCTCAGGCAGTGT 2820
Qy 3236 GGCATCAATTGCTGAGCTGTTTATTAAGGTTCAAAATTTCTGAGCTTGTGATAGGAGA 3295
Db 2821 GGCATCAATTGCTGAGCTGTTTATTAAGGTTCAAAATTTCTGAGCTTGTGATAGGAGA 2880
Qy 3296 AAAAAACAATTAGAAATTTTGAAGTTGTTGTCTGTAATGTTATGTTATGTTACCCAGAAT 3355
Db 2881 AAAAAACAATTAGAAATTTTGAAGTTGTTGTCTGTAATGTTATGTTATGTTACCCAGAAT 2940
Qy 3356 TGTGGACAGGAATTGAACAAAGGACAAAGTTTGAATTTGATTTGTAATGGCAAAAAA 3415
Db 2941 TGTGGACAGGAATTGAACAAAGGACAAAGTTTGAATTTGTAATGGCAAAAAA 3000
Qy 3416 AAAAAA 3443
Db 3001 AAAAAA 3028

RESULT 3
US-10-425-114-24930
; Sequence 24930, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24930
; LENGTH: 3589
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-221-C10_FLI
US-10-425-114-24930

Query Match 87.4%; Score 3008.6; DB 17; Length 3589;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 3226; Conservative 0; Mismatches 159; Indels 42; Gaps 8;

Qy 1 CTGGCTCGCCCTGCTCGCAATCGCGAATCTGTGAGCACCTGAGGGGTTCGAGGCCCGAG 60
Db 157 CAGCGTCGCCCTGCTTGCATCGCG-ATCGATCGATCGGAGGCCAAGAGCTAGCTAGCA 215
Qy 61 AGCTAGCTAGACGCCCGCTTCGCGCGGATGAGGCCAGCGCCGGCTGTTGGCCGG 120
Db 216 CGCGGCTCGGCTCGCGCGTGCAGCGCGATGAGGCCAGCGCGCGGCTTGTGGCCGG 275
Qy 121 CTGCGACACCGGACGAGCTGTTGCTGATCGGGGCGCACGAGGACCCAGACCGCTGCG 180
Db 276 CTGCGACACCGGACGAGCTGTTGCTCATTCGCGGCGCACGAGGACCCAGACCGCTGCG 335
Qy 181 GCGCTGAGCGGCGAGGTGTGCGAGATATGCGGCGACGAGGTTCGGGCTCAGCGTGGACGG 240
Db 336 GCGCTGAGCGGCGAGGTGTGCGAGATATGCGGCGACGAGGTTCGGGCTCAGCGTGGACGG 395
```

```
Qy 241 CGACCTCTTTCGTGCGCTGCAACGAGTGGGGTTCCTCCCGTGTGCGCGCCCTGCTACGAGTA 300
Db 396 CGACCTCTTTCGTGCGCTGCAACGAGTGGGGTTCCTCCCGTGTGCGCGCCCTGCTACGAGTA 455
Qy 301 CGAGCGCGGAGGGGACGCGAGAACTGCCCCAGTGAAGACGCGCTTACAGCGCCCTCAA 360
Db 456 CGAGCGCGGAGGGGACGCGAGAACTGCCCCAGTGAAGACGCGCTTACAGCGCCCTCAA 515
Qy 361 GGGAGCCCCGAGGGTTCCCGGGGACGATGACGAGGAGGACATCGACCTTGGAGCGACA 420
Db 516 GGGAGCCCCGAGGGTTCGCGGGGACGATGACGAGGAGGACATCGACCTTGGAGCGACA 575
Qy 421 GTTCAACATCGACGACGAGAACTCAGCAGAGGAGCTGGAGGGGACATCGAGACAGCCA 480
Db 576 GTTCAACATCGACGACGAGAACTCAGCAGAGGAGGACATCGAGACAGCCA 632
Qy 481 GATCACCGAGCGATGCTGCACGGACGATGAGTACGCGAGGGGCCCCGACGAGCGCA 540
Db 633 GATCACCGAGCGGCTGCTGCACGGACGATGAGTACGCGAGGGGCCCCGACGAGCGCA 692
Qy 541 CGGCAACAACACCCCGAGATCCCGCCATCATCACCGGTTCCCGTCCGTTCCGTTCCGTTGAG 600
Db 693 GGGCAACAACACCCCGAGATCCCGCCATCATCACCGGTTCCCGTCCGTTCCGTTCCGTTGAG 752
Qy 601 CGGTGAGTTTCCGATTAACCAACGGGTATGGCAGCGGAGGTCTCGTCTTCCCTGCACAA 660
Db 753 CGGTGAGTTTCCGATTAACCAACGGGTATGGCAGCGGAGGTCTCGTCTTCCCTGCACAA 812
Qy 661 GCGCATCATCCGTACCTGCTGCTGAGCCAGGGAGTGCCAAAGTGGGACGAGAGAAAGA 720
Db 813 GCGCATCATCCGTACCTGCTGCTGAGCCAGGGAGGCCAAGTGGGACGAGAGAAAGA 872
Qy 721 AGTGAGCTGGAAGGAGAGGATGGAACGATGGAAGTCCAAGCAGGGCATCTCT-----CGG 774
Db 873 AGTGAGCTGGAAGGAGAGGATGGAACGATGGAAGTCCAAGCAGGGCATCTCTCGGCGCGG 932
Qy 775 CGGCGGCGCGATCCCGAAGACATGAGCGCGGAGTGGCATGACGAGGAGGAGGCA 834
Db 933 CGGCGGCGGCGATCCCGAAGACATGAGCGCGGAGTGGCATGACGAGGAGGAGGCA 992
Qy 835 GCGCTCTCGAGGAAGGTGTGATCGGTCGAGCAGGTGAACCGGTACCGGATGGTGAT 894
Db 993 GCGCTCTCGGCAAGGTGTGATCGGTCGAGCAGGTGAACCGGTACCGGATGGTGAT 1052
Qy 895 CGTGCTGCTGCTCGTGTGCTCGCTTCTTCTTCTCGGTACCGTATCTCGACCCCGTCCC 954
Db 1053 CGTGCTGCTGCTCGTGTGCTCGCTTCTTCTTCTCGGTACCGTATCTCGACCCCGTCCC 1112
Qy 955 GGAGCCCATCGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGGTTCCGCATCTCGTG 1014
Db 1113 GGAGCCCATCGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGGTTCCGCCTGCTCGT 1172
Qy 1015 GATCTCGACAGTTCCCAAGTGGTTCCCATCGAGCGGAGAGTACCTTCGACCGCT 1074
Db 1173 GATCTCGACAGTTCCCAAGTGGTTCCCATCGAGCGGAGAGTACCTTCGACCGCT 1232
Qy 1075 CTCCTCTCAGGTACGAGAGGGAAGGGAGCGGTCTGCTGCTGTCGCGGTTGACCTGTTCTG 1134
Db 1233 CACCTCTCAGGTACGAGAGGGAAGGGAGCGGTCTGCTGCTGTCGCGGTTGACCTGTTCTG 1292
Qy 1135 GAGCAGGTTGACCCCGTCAAGGAGCGCGCTGTTGAGCCGCCAACCCTGCTCTCAT 1184
Db 1293 CAGCAGGTTGACCCCGTCAAGGAGCGCGCTGTTGAGCCGCCAACCCTGCTCTCAT 1352
Qy 1195 CTTGCGGCTAGACTTACCCCGTGGAGAGGTCTCTGCTAGCTTCCGACGAGCGCGGCTC 1254
Db 1353 CTTGCGGCTAGACTTACCCCGTGGAGAGGTCTCTGCTAGCTTCCGACGAGCGCGGCTC 1412
Qy 1255 GATGCTGACGTTTCAGGTTCGTCGAGAGCGGCGGAGTTTCGCGCGAAGTGGGTGCCCTT 1314
Db 1413 GATGCTGACGTTTCAGGTTCGTCGAGAGCGGCGGAGTTTCGCGCGAAGTGGGTGCCCTT 1472
Qy 1315 CTGCAAGAGTTTCGTCGATTCGAGCGCCCGCGCGGAGTTTCTATTTCTGCTCAAGGTCGA 1374
```

Db	1473	CTGCAAGAGTTCTGCATCGAGCCCGCCCGGAGTTCTATTTCTCGTCAAGGTCGA	1532
Qy	1375	CTACTCTCAAGGACAAGGTGCAGGCCACCTTCGTGCAGAGCGCGGCCCATGAAGAGAGA	1434
Db	1533	TTACTCTCAAGGACAAGGTGCAGGCCACCTTCGTGCAGAGCGCGGCCCATGAAGAGAGA	1592
Qy	1435	GTATGAGGAGTTCAAGTTCGGATCAACCGGCTGGTGGCCAAAGCCATGAAGGTGCCGGC	1494
Db	1593	GTACGAGGAGTTCAAGGTCCGCATCAACGCGCTGTGTCGCCAAGGCCATGAAGGTGCCGGC	1652
Qy	1495	AGAGGGGTGATCATATGAAGAGCGCACGCGGTGCGCCCGGGAAACAACCCCGCGACACCC	1554
Db	1653	AGAGGGGTGATCATATGAAGAGCGCACGCGGTGCGCCCGGGAAACAACCCCGCGACACCC	1712
Qy	1555	CGGCATCATCCAGGTGTTCTTGCGCACAGCGGGCGGCACGACACCCGAGGGCAACGAGCT	1614
Db	1713	TGGCATGATCCAGGTGTTCTTGCGCACAGCGGGCGGCACGACACCCGAGGGCAACGAGCT	1772
Qy	1615	GCCCGGCTCGTGTAACGTCTCCGTGAGAAAGCGCCCGGGATTCAGACACCAAGAAGGC	1674
Db	1773	GCCCGGCTCGTGTAACGTCTCCGTGAGAAAGCGCCCGGGATTCAGACACCAAGAAGGC	1832
Qy	1675	CGCGCCCATGAAACGCTCTGATTCGCGTCTCCCGCGTCTGACCAACCGGCGCATTCATGCT	1734
Db	1833	CGCGCCCATGAAACGCTCTGATTCGCGTCTCCCGCGTCTGACCAACCGGCGCATTCATGCT	1892
Qy	1735	CAACTTGGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGAGCCATGTGCTT	1794
Db	1893	CAACTTGGACTGTGATCACTACATCAACACAGCAAGGCCATCCGGAGAGCCATGTGCTT	1952
Qy	1795	CCTCATGGAACCTCAGGTCGGCCGGAAGGTCGTAGCTTCAGTTCCGTCGGCAGAGGTTCGA	1854
Db	1953	CCTCATGGAACCTCAGGTCGGCCGGAAGGTCGTAGCTTCAGTTCCGTCGGCAGAGGTTCGA	2012
Qy	1855	CGGCATCGAAGTGCACGACCGATACGCTAACAGGAAACCGCTCTTTCGACATCAACAT	1914
Db	2013	CGGCATCGAAGTGCACGACCGATACGCTAACAGGAAACCGCTCTTTCGACATCAACAT	2072
Qy	1915	GAAGGGCTCGAACGGCATCCAAAGCCGCGGTAGCTCGGACAGAGGTGCGTGTCCGGCG	1974
Db	2073	GAAGGGCTCGAACGGCATCCAAAGCCGCGGTAGCTCGGACAGAGGTGCGTGTCCGGCG	2132
Qy	1975	CAAGGCGCTCTACGGCTACAACCTCCCAAGGGACCCAAAGAGGCCCAAGATGTGACCTG	2034
Db	2133	CAAGGCGCTCTACGGCTACAACCTCCCAAGGGACCCCAAGAGGCCCAAGATGTGACCTG	2192
Qy	2035	CGACTGTCGCCGTCGTCGCGCCGCAAGAACGGAAACACGCCAAGGACGGGCTGCCGGA	2094
Db	2193	CGACTGTCGCCGTCGTCGCGCCGCAAGAACGGAAACACGCCAAGGATGGGCTGCCGGA	2252
Qy	2095	GGGCACCGCTGATATGGAGTAGATAGCGACMAAGGATGCTCATGTGCCACATGAACCTT	2154
Db	2253	GGGCACCGCTGATATGGAGTAGATAGCGACMAAGGATGCTCATGTGCCACATGAACCTT	2309
Qy	2155	CGAAGACGGTTCCGGCAGTCCGCGGCGTTTCGTACAGCTCGACGCTGATGAGAGAGCGG	2214
Db	2310	CGAAGACGGTTCCGGCAGTCCGCGGCGTTTCGTACAGCTCGACGCTGATGAGAGAGCGG	2369
Qy	2215	CGTCCCTCTTCGTCGAGCCCGCGCGCTCCTCAAGAGGGCCATCCATGTCTCAGCTG	2274
Db	2370	CGTCCCTCTTCGTCGAGCCCGCGCGCTCCTCAAGAGGGCCATCCATGTCTCAGCTG	2429
Qy	2275	CGGCTACAGGACAAGACCGACTGGGGGCTGGAGCTGGGGTGGATCTACCGGTCGATCAC	2334
Db	2430	CGGCTACAGGACAAGACCGACTGGGGGCTGGAGCTGGGGTGGATCTACCGGTCGATCAC	2489
Qy	2335	GGAGGACATCTCTGACGGGGTTCAAGATGCACTCCGCGGGTGGCGCTCCGCTGACTGCAT	2394
Db	2490	GGAGGACATCTCTGACGGGGTTCAAGATGCACTCCGCGGGTGGCGCTCCGCTGACTGCAT	2549
Qy	2395	GCCGAAGCGGCGCGTTCGAAGGGGTCGGCGCGGATCAATCTATCGGACCGCTCTCAACCA	2454

Db	2550	GCCGAAAGCGGCGCGGTTCAAGGGGTGCGCGCGGATCAACTGTTCGACGCTTCAACCA	2609
Qy	2455	GGTGCTCCGGTGGGCGCTGGGGTCCGTCGAGATCTTTCTTCAGCGGGCAAGCCCCCTGCT	2514
Db	2610	GGTGCTCCGGTGGGCGCTGGGGTCCGTCGAGATCTTTCTTCAGCGGGCAAGCCCCCTGCT	2669
Qy	2515	GTACGGCTACAAGAAAGGCAACTCAAGTGGCTGGAGCGTTCCGCTACATCAACACCA	2574
Db	2670	GTACGGCTACAAGAAAGGCAACTCAAGTGGCTGGAGCGTTCCGCTACATCAACACCA	2729
Qy	2575	CATCTACCCCTTCACTCCGTCGCGTCCGCTCGCTACTGTACCTCCGCGCTGCGCT	2634
Db	2730	CATCTACCCCTTCACTCCGTCGCGTCCGCTCGCTACTGTACCTCCGCGCTGCGCT	2789
Qy	2635	CCTCACGGCAAGTTTCATATGCGCTCGATTAGCACGTTCCCGCAGCCTCTTCTTCATCGC	2694
Db	2790	CCTCACGGCAAGTTTCATATGCGCTCGATTAGCACGTTCCCGCAGCCTCTTCTTCATCGC	2849
Qy	2695	CCTCTTCATGTCCATCTTCCGGACGGGCATCTTGGAGATCGGTGGAGCGGGGTGAGCAT	2754
Db	2850	CCTCTTCATGTCCATCTTCCGGACGGGCATCTTGGAGATCGGTGGAGCGGGGTGAGCAT	2909
Qy	2755	CGAGGAGTGTGGAGAAACGAGCAGTCTTGCGTCAATCGCGCGGTGTCCCGCATCTCTT	2814
Db	2910	CGAGGAGTGTGGAGAAACGAGCAGTCTTGCGTCAATCGCGCGGTGTCCCGCATCTCTT	2969
Qy	2815	CGCGTCTGTGAGGGCTGTCTCAAGTCTCTCGCGGGATCGACACCAATCTCACCGTCAC	2874
Db	2970	CGCGTCTGTGAGGGCTGTCTCAAGTCTCTCGCGGGATCGACACCAATCTCACCGTCAC	3029
Qy	2875	CTCCAGGCGCACGGGCGAGGACGACGAGTTTCGCGAGCTCTACGCCCTTCAAGTGCAC	2934
Db	3030	CTCCAGGCGCACGGGCGAGTGTGACGACGAGTTTCGCGAGCTCTACGCCCTTCAAGTGCAC	3089
Qy	2935	CACGCTCCTCATCCCGCCCAACCGTGTCTCATCTTAACGTTCATCGGCGTCTGGCGCGG	2994
Db	3090	CACGCTCCTCATCCCGCCCAACCGTGTCTCATCTTAACGTTCATCGGCGTCTGGCGCGG	3149
Qy	2995	CATCTCCGACGCCATCAACAAACGGGTACAGTCTTGGGGGCCCTCTTCGGCAAGCTCTT	3054
Db	3150	CATCTCCGACGCCATCAACAAACGGGTACAGTCTTGGGGGCCCTCTTCGGCAAGCTCTT	3209
Qy	3055	CTTCGCTTCTGGGTCACTCGTCAACCTCTACCCGTTCTCCTCAAGGGGCTCATGGGCGCCA	3114
Db	3210	CTTCGCTTCTGGGTCACTCGTCAACCTCTACCCGTTCTCCTCAAGGGGCTCATGGGCGCCA	3269
Qy	3115	GAAACAGCGCCACCGTTGTGTTCATCTGGTTCATCTGCTGGCTTCACATCTTCTCCCT	3174
Db	3270	GAAACAGCGCCACCGTTGTGTTCATCTGGTTCATCTGCTGGCTTCACATCTTCTCCCT	3329
Qy	3175	GCTCTGGGTCAAGGATCGACCCCTTTCATCGTCAGACCAAGGGCCCGACGCTCAGGCAAGT	3234
Db	3330	GCTCTGGGTCAAGGATCGATCCCTTCATCTGTCAGACCAAGGGCCCGACGCTCAGGCAAGT	3389
Qy	3235	TGGCATCAATGTGTAGCTGTTTATTAAAGTTCAAAA---TTCGGAGCTTGTGCATAGG	3294
Db	3390	TGGCATCAACTGTGTAGCTGCCTATCACGGTTCAAAAAAGTTCTAGAGCTTGTGCATAGG	3449
Qy	3292	GAGAAAAAACATTTAGAAATTTTGT-----AGGTTGTTGTG	3350
Db	3450	GAG---AAACAAACAGTTTAAAGAAATTTGTGTTGTGATGTTTGGAGCTCATATGTCATGTTT	3508
Qy	3331	TCTGTAATGTTATGGTATCCAGAAATTTGT---CGGACGAGGAATTTGAAACAAAGCAAGG	3386
Db	3509	TGTGTAATGTTCTGGACCCAGAAATTTGTCCGCGCGACGAGGAATTTGAAACAAAGCAAGG	3568
Qy	3387	TTTGATTT 3393	
Db	3569	TTTGAAT 3575	

; Sequence 15093, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15093
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20969C.1
US-10-437-963-15093

Query Match 78.2%; Score 2692.8; DB 18; Length 3448;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2954; Conservative 0; Mismatches 322; Indels 18; Gaps 5;

QY 67 CCTAGCAGCGCGCTCCGCGCGATGGAGGCCAGCGCGGCTGGTGGCGGCTCGCA 126
DB 3398 CGTTGCGCGCGCTGGTGGCGCGCATGGAGCGCGCGCGGCTGGTGGCGGCTCGCA 3339
QY 127 CAACCGGAACGAGCTGTGTGATCCGGGGCCAGAGGACCCCAAGCGCTGCGGGGCT 186
DB 3338 CAACCGGAACGAGCTGTGTGATCCGGGGCCAGAGGACCCCAAGCGCTGCGGGGCT 3279
QY 187 GAGCGGGAGGTGTGCGGAGATATCGCGGACGAGGTGGGTCTACGGTGGACGCGACCT 246
DB 3278 GAGCGGGAGGTGTGCGGAGATATCGCGGACGAGGTGGGTCTACGGTGGACGCGACCT 3219
QY 247 CTTGCTGCTGCAACGAGTGGCTTCCCGTGTGCGGCGCTGCTACGAGTACGAGCG 306
DB 3218 CTTGCTGCTGCAACGAGTGGCTTCCCGTGTGCGGCGCTGCTACGAGTACGAGCG 3159
QY 307 CCGGAGGGGACGCGAGAACTGCCCGCAGTGAAGAGCGCTTACAGCGCTCAAGGGGAG 366
DB 3158 CCGGAGGGGACGCGAGAACTGCCCGCAGTGAAGAGCGCTTACAGCGCTCAAGGGGAG 3099
QY 367 CCCGAGGGTTCGCGGGACGATGACGAGGAGGATCGACGACCTGAGGACGAGTTCAA 426
DB 3098 CCCGAGGGTTCGCGGGACGATGACGAGGAGGATGACGACCTGAGGACGAGTTCAA 3039
QY 427 CATCGACGACGAGATCAGCAGAGGACGAGTGGGCGAA-----CATGACAGACGCA 480
DB 3038 CATCGACGACGAGATCAGCAGAGGACGAGTGGGCGAA-----CATGACAGACGCA 2979
QY 481 GATCAGGAGGAGTGTGTCAGCGGACGAGTACGAGGAGGGGCGCGACGAGCGGA 540
DB 2978 CATCAGGAGGAGTGTGTCAGCGGACGAGTACGAGGAGGGGCGCGACGAGCGGA 2919
QY 541 CGGCAACAAACCCCGGAGATCCCGGCGATCATCACCGGCTCCCGTCCGTCGCGGTGAG 600
DB 2918 CGGCAACAAACCCCGGAGTCCCGGCGATCATCACCGGCTCCCGTCCGTCGCGGTGAG 2862
QY 601 CGGTGAGTTCCGATTACCAACGGTATGGCCAGCGGAGTCTCGTCTTCCCTGCAAA 660
DB 2861 CGGTGAGTTCCGATTACCAACGGTATGGCCAGCGGAGTCTCGTCTTCCCTGCAAA 2802
QY 661 GCGCATCCATCCGTACCTGTGTGAGCAGGAGGTGCCAAGTGGGACGAGAGAAAGA 720
DB 2801 GCGCATCCATCCGTACCTGTGTGAGCAGGAGGTGCCAAGTGGGACGAGAGAAAGA 2742

QY 721 AGTCAGCTGGAAGAGGAGGATGGACGATGGAAGTCCAAGCAGGGGCATCTCTCGCGCGG 780
DB 2741 GGTGAGCTGGAAGAGGAGGATGGACGATGGAAGTCCAAGCAGGGGCATCTCTCGCGCGG 2682
QY 781 CG----CGATCCCGAAGACATGAGCGCGGAGCTGGCACTGAACACGAGGGGAGCGGCC 837
DB 2681 CGCCCCGATCCCGAAGACATGAGCGCGGAGCTGGCACTGAACACGAGGGGAGCGGCC 2632
QY 838 GCTGTCGAGGAAGGTGTCCGATCGGATCGAGCAAGGTGAACCGGTACCGGATGGTGCAT 897
DB 2621 GCTGTCGAGGAAGGTGTCCGATCGGATCGAGCAAGGTGAACCGGTACCGGATGGTGCAT 2562
QY 898 GGTGCGTCTGTTGTCGCTTCTTCTCCCGTACCGTATCTGTCGACCCCGTCCCGGA 957
DB 2561 CTTCCGCTCTGTTGTCGCTTCTTCTCCCGTACCGTATCTGTCGACCCCGTCCCGGA 2502
QY 958 CGCATCGGCTGTGGCTCGTCTCCATCATCTCGGAGATCTGGTTCGCCATCTCTCTGAT 1017
DB 2501 CGCATCGGCTGTGGCTCGTCTCCATCATCTCGGAGATCTGGTTCGCCATCTCTCTGAT 2442
QY 1018 CTTGACACAGTTTCCCAAGTGGTTCCTCATCGACCGGAGACGTACCTCGACCGCTCTC 1077
DB 2441 CTTGACACAGTTTCCCAAGTGGTTCCTCATCGACCGGAGACGTACCTCGACCGCTCTC 2382
QY 1078 CTTGACGATCGAGGAGGAGCGGTGCTGCTGTCGCGGTGACCTGTTGTCGAG 1137
DB 2381 CTTGCGCTACGAGCGCGGAGGAGCGGTGCTGCTGTCGCGGTGACCTGTTGTCGAG 2322
QY 1138 CAGGTCGACCGCTCAAGGAGCGCGCTGTCGACCGCAACACCGTCTCTCATCTCT 1197
DB 2321 CAGGTCGATCTCAAGGAGCGCGCTGTCGACCGCAACACCGTCTCTCATCTCT 2262
QY 1198 CGCGTGAAGTACCTCCCGTGAAGTCTCTGTCGACCGTCTCGACCGCGGTGCT 1257
DB 2261 CGCGTGAAGTACCTCCCGTGAAGTCTCTGTCGACCGTCTCGACCGCGGTGCT 2202
QY 1258 GCTGACGTTTCAAGTCTGTCGAGAGCGCGGAGTTGTCGCGGAGTGGTGGCTTCTG 1317
DB 2201 GCTGACGTTTCAAGTCTGTCGAGAGCGCGGAGTTGTCGCGGAGTGGTGGCTTCTG 2142
QY 1318 CAAGAGTTCGCGCATCGAGCGCGCGCGCGGAGTTCTACTTCTCGCTCAAGTCTGACTA 1377
DB 2141 CAAGAGTTCGCGCATCGAGCGCGCGCGCGGAGTTCTACTTCTCGCGAGGTGCGACTA 2082
QY 1378 CTTCAAGGAACAAGTGTGAGCGCTTCTGTCGAGGAGCGCGCGGATGAAGAGAGTA 1437
DB 2081 CTTCAAGGAACAAGTGTGAGCGCTTCTGTCGAGGAGCGCGCGGATGAAGAGAGTA 2022
QY 1438 TGAGGAGTTCAAGTTCGAGTCAACGCGCTGGTGGCCAAAGGCGATGAAGTGGCGGAGA 1497
DB 2021 CGAGGAGTTCAAGTTCGAGTCAACGCGCTGGTGGCGGAGGCGGAGGAGTGGCGGAGA 1962
QY 1498 GGGGTGATCATCAAGGAGCGCGCGCGCGCGGAGCAACACCGCGGACCAACCGCGG 1557
DB 1961 AGGTTGATCATGAAGGAGCGCGCGCGCGCGGAGCAACACCGCGGACCAACCGCGG 1902
QY 1558 CATGATCCAGTGTCTTCTGGGCGACGCGCGCGCGCGGAGCAACACCGAGGTGCG 1617
DB 1901 CATGATCCAGTGTCTTCTGGGCGACGCGCGCGCGCGGAGCAACACCGAGGTGCG 1842
QY 1618 CGCGCTCGTGTGATCGTCTCGGTGAGAGCGCGCGGATTCAGCAACCAAGAGGCGCG 1677
DB 1841 CGCGCTCGTGTGATCGTCTCGGTGAGAGCGCGCGCGGATTCAGCAACCAAGAGGCGCG 1782
QY 1678 CGCATGAGCGCTGATGTCGCTCTCGCGGTGTCGACCAACCGCGGATTCATGCTCAA 1737
DB 1781 CGCATGAGCGCGCTGATGTCGCTCTCGCGGTGTCGACCAACCGCGGATTCATGCTCAA 1722
QY 1738 CTTGAGTGTGATGATCATCAATCAACAGAGGAGGATCCGCGGAGGCGCATGTCCTCT 1797
DB 1721 CTTGAGTGTGATGATCATCAATCAACAGAGGAGGATCCGCGGAGGCGCATGTCCTCT 1662
QY 1798 CATGGAACCTTCAGGTTCGCGCGGAGGATGTCGTAAGTTCCTCGGAGAGGTTCGACGG 1857

Db 1661 CATGGATCCGAGTCCGACGGAAGGTTTGCTAGCTGCAGTTCCCGAGAGGTTTCGACGG 1602
Qy 1858 CATCGAGTGCACGACGATACGCTAAACAGGAACACCGTCTCTTCTTCGACATCAACATGAA 1917
Db 1601 CATCGAGTGCACGACGATACGCTAAACAGGAACACCGTCTCTTCTTCGACATCAACATGAA 1542
Qy 1918 GGGCTCGAGCGCATCCAAAGCCCGGTGTAGCTGCGGACAGGGTGCCTGTTCCGCGGCCA 1977
Db 1541 GGGCTTGTATGGATCCAGGCGCCGGTGTAGCTGCGGACAGGGTGCCTGTTCCAGCGGCA 1482
Qy 1978 GGGCTTACGGCTACACCTCCCAAGGACCCCAAGAGGCCCAAGATGATGATCGGAC 2037
Db 1481 GGGCTGTACGATACCAACCCACCAAGGACCCCAAGAGGCCCAAGATGATGATCGGAC 1422
Qy 2038 CTGCTGCCCGTCTTCGGCCCAAGAGCGGAAACACGCCAAGAGCGGGCTCCCGGAG-- 2095
Db 1421 CTGCTGCCCGTCTTCGGGAGGAAGAGCGGAGCACCGGAGCGGCTCCCGGAGGC 1362
Qy 2096 -GGCAACCGCTGATATGGAGTATAGTAGCGAACAGGAGATGCTCATGTCCACATGAACATT 2154
Db 1361 CGTCCCGCCGACCGCGGGATGGACAGCGAACAGGAGATGCTCATGTCCGAGATGAATT 1302
Qy 2155 CGAGAACGGTTCCGGCAGTCCGCGGCTTCGTCAGCTGACGCTGATGAGAGAGCGG 2214
Db 1301 CGAGAACGGTTCCGGCAGTCCGCGGCTTCGTCAGCTGACGCTGATGAGAGAGCGG 1242
Qy 2215 CGTCCCTCTTCGTCGAGCCCGCGCTCCCTCAAGAGGSCCATCATGTCATCAGCTG 2274
Db 1241 CGTCCCGCTGCTGTCAGCCCGCGCTCCCTCAAGAGGCCATCATGTCATCAGCTG 1182
Qy 2275 CGCTACGAGGACAAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTGATCAC 2334
Db 1181 CGGCTACGAGACAAAGACCGACTGGGGTCTCGAGCTGGGGTGGATCTACGGGTGATCAC 1122
Qy 2335 GGAGGACATCTGACGGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGTTGATCTGCAT 2394
Db 1121 GGAGGACATCTTAACGGGGTTCAAGATGCACTGCGCGGGTGGAGGTGCGGTACTGCAT 1062
Qy 2395 GCCAAGCGGGCGGCTTCAAGGGTTCGGCGCGGATCAATCTATCGGACCGTCTCAACCA 2454
Db 1061 GCCGAGAGGGCGGCTTCAAGGGTTCAGGGGTCAGGGCGGATCACTATCGACCGTCTCAACCA 1002
Qy 2455 GTGCTCCGGTGGCGCTGGGGTTCGTCGAGATCTTCTCAGCGGACAGCCCCCTGCT 2514
Db 1001 GTGCTCCGGTGGCGCTCGGCTCCGTCGAGATCTTCTCAGCGGACAGCGCGCTCCT 942
Qy 2515 GTACGGCTACAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCCCTACATCAACACAC 2574
Db 941 CTACGGCTACAGAACGGCAACCTCAAGTGGCTTCGAGCGCTTCTCTTACATCAACACAC 882
Qy 2575 CATCTACCCCTTCACTCCGCTCGCTCGCTACTGCACTCCCTCCCGCGCTCTGCGCT 2634
Db 881 CATCTACCCCTTCACTTCTCCTCCCTCTCGCTACTGCACTCCCGCGCTCTGCGCT 822
Qy 2635 CCTCACCGCAAGTTCATATGCGGTGATAGCAGCTTCGCCAGCGCTTCTTCTATCGC 2694
Db 821 CCTCACCGCAAGTTCATATGCGGTGATAGCAGCTTCGCCAGCTTGTCTTCTATCGC 762
Qy 2695 CCTCTTATGTCATCTTCGGAGCGGCACTTCGGAGATCGGTGGAGCGGGGTGAGCAT 2754
Db 761 GCTCTTATCTCATCTTTCGAGCGGCACTTCGGAGATGAGGTGGAGCGGGGTGAGCAT 702
Qy 2755 CGAGGAGTGTGGAGGAACAGACAGTTCTGGGTATCGGGGGGTGTCGCCGATCTCTT 2814
Db 701 CGAGGAGTGTGGAGGAACAGACAGTTCTGGGTATCGGGGGGTGTCGCCGATCTCTT 642
Qy 2815 CGCGCTGTGACAGGGCTGTCTAAGGTCTTCGCGGGATCGACACCAATTCACCGCTCAC 2874
Db 641 CGCGGTGTGACAGGGCTGTCTAAGGTGTCTGCGGGATCGACACCAATTCACCGCTCAC 582
Qy 2875 CTCCAGGCCACCGGCGACAGAGACGAGTTTCGCCGAGCTCTACGCCCTTCAAGTGGAC 2934

Db 581 GTCAGAGGCCACCGGAGACGAGGACGAGCTTCGCGAGCTCTACGCCCTTCAAGTGGAC 522
Qy 2935 CAGCTCTCTCATCCCGCCACACCGTCTCATCATTAAGCTCATCGCGCTCTGTCGCGCGG 2994
Db 521 CACCTCTCTCATCCCGCCACACCGTCTCATCTCAACATCATCGCGCTCTGTCGCGCGG 462
Qy 2995 CATCTCCGACGCGCATCAACACGCGGTACAGTCTCTGGGGGCCCTCTTTCGCGCAAGCTCTT 3054
Db 461 CGTCTCCGACGCGCATCAACACGCGTCCGAGGCGTGGGGCCCGCTCTTTCGCGAAGCTCTT 402
Qy 3055 CTTTCGCTTCTGGGTTCATCGTCCACCTCTACCGTTCCTCAAGGGGCTCATGGGGCGGCCA 3114
Db 401 CTTTCGCTTCTGGGTTCATCGTCCACCTCTACCGTTCCTCAAGGGGCTCATGGGGAGGCA 342
Qy 3115 GAACAGAGCGCCACCGTTCGTCATCTGGTCCATCTGCTGGCTCCATCTTCTCCCT 3174
Db 341 GAACCGGACGCCACCATTTGTCATCTGGTCCGTCTGCTCGCTCCATCTTCTTCTCTT 282
Qy 3175 GCTCTGGGTACGATCGACCCCTTTTCATCGTCAGGACCAAGGGCCCGGACGCTCAGGCGATG 3234
Db 281 GCTCTGGGTACGATTCCTTTCACCATCAAGGCCAGGGGCCCTGACGTCAGGCGATG 222
Qy 3235 TGGCATCAATTGCTGAGCTGTTTATTAAGGTTCAAAATTCGAGCTTGTGCTATAGGAG 3294
Db 221 CGGCATCAACTGCTGAGAGAGGGCGTCAG--CGATTTCTGAAGATTTGTGCTATAGGGG 165
Qy 3295 AAAAAACAATTTAGAAATTTTGAAGTTGTTGCTCTGTAATGTTATGCTAC 3348
Db 164 GCAGCAAGAAGATCGAATTTGTAATTTGTTGTTGCTGTTGAGTTCTGTC 111

RESULT 5

US-10-425-114-833

; Sequence 833, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jindong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 833

; LENGTH: 2687

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700093978_FLI

US-10-425-114-833

Query Match

Best Local Similarity 76.5%; Score 2633.2; DB 17; Length 2687;

Matches 2647; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 773 GCGCGCGCGCGCATCCCGAAGAGTTCGATCGCGTCGAGCAAGTGAACCCGCTACCGATGGTG 832
Db 1 GCGCGCGCGCGCATCCCGAAGAGTTCGATCGCGTCGAGCACTGAACGACGAGCGAGG 60
Qy 833 CAGCCCTGTGAGGAAGTTCGATCGCGTCGAGCAAGTGAACCCGCTACCGATGGTG 892
Db 61 CAGCCCTGTGAGGAAGTTCGATCGCGTCGAGCAAGTGAACCCGCTACCGATGGTG 120
Qy 893 ATCGTGGTGGTCTCGTTCGCTTCCTCCGCTACCGTATCTTCCGACCCGCTC 952
Db 121 ATCGTGGTGGTCTCGTTCGCTTCCTCCGCTACCGTATCTTCCGACCCGCTC 180
Qy 953 CCGGACGCCATCGGGCTGTGGTCTGCTCCATCATCTGCGAGATCTGGTTCGCCATCTCC 1012


```
QY 3173 CTGCTCTGGGTGAGGATCGACCCCTTTCATCGTCAGACCAAGGCGCCGACGTCAGGCAG 3232
Db 2401 CTGCTCTGGGTGAGGATCGACCCCTTTCATCGTCAGACCAAGGCGCCGACGTCAGGCAG 2460
QY 3233 TGTGGCATCAATTTGCTGAGCTGTTTATTAGGTTCAAAATTCGGAGCTTGTGCATAGG 3292
Db 2461 TGTGGCATCAATTTGCTGAGCTGTTTATTAGGTTCAAAATTCGGAGCTTGTGCATAGG 2520
QY 3293 AGAAAAAACAATTTAGAAATTTTGAAGTTGTTGTCTGTAATGTTATGTAACCCAG 3352
Db 2521 AGAAAAAACAATTTAGAAATTTTGAAGTTGTTGTCTGTAATGTTATGTAACCCAG 2580
QY 3353 AATTGTCGACGAGGAATGAACAAAGGACAAGGTTTGATTTGTAATGSCAAAAA 3412
Db 2581 AATTGTCGACGAGGAATGAACAAAGGACAAGGTTTGATTTGTAATGSCAAATGAGTG 2640
QY 3413 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3442
Db 2641 CAGTCATGCTAAATGTAATCCAAATAGCAA 2670

RESULT 6
US-10-425-115-139789
; Sequence 139789, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 139789
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_58974C.1
US-10-425-115-139789

Query Match 72.5%; Score 2495.8; DB 18; Length 2830;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2653; Conservative 0; Mismatches 102; Indels 36; Gaps 6;

QY 637 CGAGGTCTCGTCTCCCTGCACAAGCGCATCCATCCCTGCTGTCTGAGCCAGGAG 696
Db 1 CGAGCTCTCGTCTCCCTGCACAAGCGCATCCATCCCTGCTGTCTGAGCCAGGAG 60
QY 697 TGCCAAGTGGGACGAGAGAAAGTGAAGTGAAGTGAAGGAGAGATGAGACTGGAATC 756
Db 61 CGCAAAGTGGGACGAGAGAAAG-AGTGAAGTGGAAAGGAGAGATGAGACTGGAATC 119
QY 757 CAACGAGGGCATCTCTCGCGCGCGC-----GCCGATCCGGAAGACATGAGCCGCACGT 810
Db 120 CAAGCAGGGCATCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
QY 811 GGCATCGAACGACGAGCGAGGCGCGCTGTGTCGAGGAAGGTGTCGATCGCGTCGAGCAA 870
Db 180 GCGCGTGAACGACGAGCGAGGCGCGCTGTGTCGAGGAAGGTGTCGATCGCGTCGAGCAA 239
QY 871 GGTGAACCGGTACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930
Db 240 GGTGAACCGGTACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299
QY 931 GTACCGTATCTTCACCCCGCTCCCGGACGCCATCGGCGTGTGGCTGTCTCCATCATCTG 990
Db 300 CTACCGCATCTTCACCCCTCTCCGAGCGCATCGGCGTGTGGCTGTCTCCATCATCTG 359
```

```
QY 991 CGAGATCTGCTGCGCCATCTCTGAGATCTCTGACCAAGTTCCTCCCAATGTTCCCATCGA 1050
Db 360 CGAGATCTGCTGCGCGTCTCTGAGTCTCTGACCAAGTTCCTCCCAATGTTCCCATCGA 419
QY 1051 CCGGAGAGCTACTCTGACCGGCTCTCTCAGGTACGAGAGGAGAGGAGCGGTGCGT 1110
Db 420 CCGGAGAGCTACTCTGACCGGCTCTCTCAGGTACGAGAGGAGAGGAGCGGTGCGT 479
QY 1111 GCTGTGCGCGGTGAGACCTGTTCTGAGCAGCGGTGAGACCGCTCAAGGAGCGCGGTGCT 1170
Db 480 GCTGTGCGCGGTGAGACCTGTTCTGAGCAGCGGTGAGACCGCTCAAGGAGCGCGGTGCT 539
QY 1171 GACCGCAACAACGCTGCTCTCATCTCTGCGGTGAGTACCTCCGCTGGAACAAGTCTCTCTG 1230
Db 540 CACCGCAACAACGCTGCTCTCATCTCTGCGGTGAGTACCTCCGCTGGAACAAGTCTCTCTG 599
QY 1231 CTACGCTCTCCGAGCAGCGCGCTGATGCTGAGTCTGAGTCTGAGTCTGAGAGCGGCCGA 1290
Db 600 CTACGCTCTCCGAGCAGCGCGCTGATGCTCAGTCTGAGGCGCTGTCGAGAGCGGCCGA 659
QY 1291 GTTTCGCGCAAGTGGTGCCTTCTGCAAGAGTTCGGGCATCGAGCCCCCGCGCCCGGA 1350
Db 660 GTTTCGCGCAAGTGGTGCCTTCTGCAAGAGTTCGATCGAGCCCCCGCGCCCGGA 719
QY 1351 GTTCTACTTCTCGCTCAAGGTGCACTACTCAAGGACAAGGTGTCAGCCACTCTGTCGA 1410
Db 720 GTTCTACTTCTCGCTCAAGGTGCACTACTCAAGGACAAGGTGTCAGCCACTCTGTCGA 779
QY 1411 GGAGCGCGCGCCATGAAGAGAGATGATGAGGAGTTCAGGTCGGATCAACGCGCTGCT 1470
Db 780 GGAGCGCGCGCCATGAAGAGAGATGATGAGGAGTTCAGGTCGGATCAACGCGCTGCT 839
QY 1471 GGCCAAGGCGCATGAAGTTCGCGCAGAGGCGTGCATCAAGAGGACGCGACGCGTGGCC 1530
Db 840 GGCCAAGGCGCATGAAGTTCGCGCAGAGGCGTGCATCAAGAGGACGCGACGCGTGGCC 899
QY 1531 CGGGAACAACACCGCGCACACCCCGCATGATCCAGGTGTTCTCTGGGCGACAGCGCGG 1590
Db 900 CGGGAACAACACCGCGCACACCCCTGGCATGATCCAGGTGTTCTCTGGGCGACAGCGCGG 959
QY 1591 CACGACACCGAGGCGCAAGGTCGCCCGCTCGGTGATGCTCTCCGTCGAGAGCGGCC 1650
Db 960 CACGACACCGAGGCGCAAGGTCGCCCGCTCGGTGATGCTCTCCGTCGAGAGCGGCC 1019
QY 1651 GGGATTCAGACCAACAAGGCGCGCCCATGAACGCTCTGATTCGGCTCTCCCGCGT 1710
Db 1020 GGGATTCAGACCAACAAGGCGCGCCCATGAACGCTCTGATTCGGCTCTCCCGCGT 1079
QY 1711 GCTGACCAACGCGCCCATCATGCTCAAATTGAGCTGTGATCACTACATCAACAACAGCAA 1770
Db 1080 GCTGACCAACGCGCCCATCATGCTCAAATTGAGCTGTGATCACTACATCAACAACAGCAA 1139
QY 1771 GGCATTCGCGGAGGCGCATGCTCTCATGACCTCGAGTCCGCGGAGGCTGCTA 1830
Db 1140 GGCATTCGCGGAGGCGCATGCTCTCATGACCTCGAGTCCGCGGAGGCTGCTA 1199
QY 1831 CGTTCAGTTCGCGCAGAGGTCGACGCGCATCGACGTCGACGATAGCTTAACAGGAA 1890
Db 1200 CGTTCAGTTCGCGCAGAGGTCGACGCGCATCGACGTCGACGATAGCTTAACAGGAA 1259
QY 1891 CACCGTCTTCTTCGACATCAACATGAAGGCGTGAACGCGATCCAGAGCCGCTGATCGT 1950
Db 1260 CACCGTCTTCTTCGACATCAACATGAAGGCGTGAACGCGATCCAGAGCCGCTGATCGT 1319
QY 1951 CGGGAAGGCGTGGTTCGCGCGCGAGGCGTCTACGCTCAACCCCTCCCAAGGAGCC 2010
Db 1320 CGGGAAGGCGTGGTTCGCGCGCGAGGCGTCTACGCTCAACCCCTCCCAAGGAGCC 1379
QY 2011 CAAGAGGCGCAAGATGCTGACCTGCGCTGCTGCGCGTCTTCCGCGCAAGAGCGGAA 2070
Db 1380 CAAGAGGCGCAAGATGCTGACCTGCGCTGCTGCGCGTCTTCCGCGCAAGAGCGGAA 1439
QY 2071 ACACGCCAAGGAGCGGCTCCCGAGGCGCACCGCTGATATGGAGTAGATAGCGACAGGA 2130
```

Db 1440 GGACGCCAAGATGGCTGCCGAGGAGCACCGCGGATATCGGATGACACGCGCAAGGA 1499
Qy |||||
Db 2131 GATGCTCATGTGCCACATGAATCTTCGAGAACGGTTCCGGCAGTCCCGCGGCTTCGTCCAC 2190
Qy |||||
Db 1500 GATGCTCATGTGCCAGATGAATCTTCGAGAACGGTTCCGGCAGTCCCGCGGCTTCGTCCAC 1559
Qy |||||
Db 2191 GTCACGCTGATGAGGAGAACGGGCTCCCTCTTCGTTCGAGCCCGCGGCTTCCTCAA 2250
Qy |||||
Db 1560 GTCACGCTCATGAGGAGAACGGGCTCCCTCTTCGTTCGAGCCCGAGCTGCGCTCTCAA 1619
Qy |||||
Db 2251 GGAGCCATCATGTCTCATCTCGGCTACGAGCAAGACCGACCTGGGGCTGGAGCT 2310
Qy |||||
Db 1620 GGAGCCATCATCTCATCTCGGCTACGAGCAAGACCGACCTGGGGCTGGAGCT 1679
Qy |||||
Db 2311 GGGGTGATCTACGGGTGATACGAGGAGCATCTGACGGGGTTCAAGATGCACTGCCG 2370
Qy |||||
Db 1680 GGGGTGATCTACGGGTGATACGAGGAGCATCTGACGGGGTTCAAGATGCACTGCCG 1739
Qy |||||
Db 2371 CGGTTGGCTCCGTGTACTGATCGAGCGAGCGGGCGGCTTCAAGGGGTCCGGCGCGAT 2430
Qy |||||
Db 1740 CGGTTGGCTCCGTGTACTGATCGAGCGAGCGGGCGGCTTCAAGGGGTCCGGCGCGAT 1799
Qy |||||
Db 2431 CAATCTATCGGACCGTCTCAACCAAGGTGCTCCGGTGGGCGCTGGGTCCGTCGAGATCTT 2490
Qy |||||
Db 1800 CAACCTGTCCGACCGTCTCAACCAAGGTGCTCCGGTGGGCGCTGGGTCCGTCGAGATCTT 1859
Qy |||||
Db 2491 CTTGAGCGGACACCGCCCTCTGTGTACGGTACAGAAACCGCAACCTCAAGTCCGCTGA 2550
Qy |||||
Db 1860 CTTGAGCGGACACCGCCCTCTGTGTACGGTACAGAAACCGCAACCTCAAGTCCGCTGA 1919
Qy |||||
Db 2551 GCGTTGCTTACATCAACACCACTATCCCTTCACTCGCTCCGCTGCTGCGCTA 2610
Qy |||||
Db 1920 GCGTTGCTTACATCAACACCACTATCCCTTCACTCGCTCCGCTGCTGCGCTA 1979
Qy |||||
Db 2611 CTGACACCTCCCGCGTCTGCTCCCTCACCGGCAAGTTCACTATGCGCTGATAGCAC 2670
Qy |||||
Db 1980 CTGTACCTGCGACCGTCTGCTCCCTCACCGGCAAGTTCACTATGCGCTGATCAGCAC 2039
Qy |||||
Db 2671 GTTCGCGACGCTCTTCTTCATCGCCCTCTTCACTGCTCATCTTCGCGACGGGCACTCTGGA 2730
Qy |||||
Db 2040 GTTCGCGACGCTCTTCTTCATCGCGCTCTTCACTGCTCATCTTCGCGACGGGCACTCTGGA 2099
Qy |||||
Db 2731 GATCGGTGAGCGGGTGAGCATCGAGAGTGTGAGAGAAACGAGCAGTTCTGGGTCTAT 2790
Qy |||||
Db 2100 GATCGGTGAGCGGGTGAGCATCGAGAGTGTGAGAGAAACGAGCAGTTCTGGGTCTAT 2159
Qy |||||
Db 2791 CGGCGGCTGTCGCGCATCTTTCGCGCTGTCAGGGCTGCTCAAGTCTCTCGCGG 2850
Qy |||||
Db 2160 CGGCGGCTGTCGCGCATCTTTCGCGCTGTCAGGGCTGCTCAAGTCTCTGGCGCG 2219
Qy |||||
Db 2851 GATCGACACCAACTTTCACCGTCACTTCAAGGCCACCGCGCAACGAGGACGACGAGTTCCG 2910
Qy |||||
Db 2220 GATCGACACCAACTTTCACCGTCACTTCAAGGCCACCGCGCAACGAGGACGAGTTCCG 2279
Qy |||||
Db 2911 CGAGCTTACCGCTTCAAGTGAACAGCTCTCTCATCCGCGCCACACGCTGCTCATCAT 2970
Qy |||||
Db 2280 CGAGCTTACCGCTTCAAGTGAACAGCTCTCTCATCCGCGCCACACGCTGCTCATCAT 2339
Qy |||||
Db 2971 TAACTCATTCGGCTGTCGGCGGATCTCGAGCCATCAACACCGGTACCACTCTCTG 3030
Qy |||||
Db 2340 CAACATCATTCGGCTGTCGGCGGATCTTCGAGCCATCAACACCGGTACCACTCTCTG 2399
Qy |||||
Db 3031 GGGGCCCTCTTTCGGCAAGCTCTTTCGCGCTTCTGGGTCTATCGTCCACCTCTACCGGTT 3090
Qy |||||
Db 2400 GGGGCCCTCTTTCGGCAAGCTCTTTCGCGCTTCTGGGTCTATCGTCCACCTCTACCGGTT 2459
Qy |||||
Db 3091 CCTCAAGGGCTCATGCGGCGCCAGAAACAGGACGCCACCGTGTGTGTCTGTCATCTGGTCCAT 3150
Qy |||||
Db 2460 CCTCAAGGGCTCATGCGGCGCCAGAAACAGGACGCCACCGTGTGTGTCTGTCATCTGGTCCGT 2519
Qy |||||
Db 3151 TCTGCTGGCTCCATCTTCTCCCTGCTCTGGGTGAGGATCGACCCCTTTCATCTGTCAGGAC 3210
Qy |||||

Db 2520 GCTGCTGGCTCCATCTTCTCCCTGCTCTCGGTCAGGATCGACCCCTTTCATCTGTCAGGAC 2579
Qy |||||
Db 3211 CAAGGGCCCGGACGTCAGGAGTGTGGCATCAATTTCTGAGCTGTATTAAAGGTTCAAA 3270
Qy |||||
Db 2580 CAAGGGCCCGGACGTCAGGAGTGTGGCATCAATTTCTGAGCTGTATTAAAGGTTCAAA 2639
Qy |||||
Db 3271 A---TTCTGAGCTTGTGCATAGGAGAAAACAAATTTAGAAATTTCTTAAGTTGTT 3327
Qy |||||
Db 2640 AAAGTTCTAGAGCTTGTGCATAGGAG-AAACAACAGTTTAGAAATTTTGTATTGTCTATG 2698
Qy |||||
Db 3328 GTGTC-----TGTAATGTTATGTTACCCAGAAATTTGT-----CGGA 3362
Qy |||||
Db 2699 TTTTGGAGCTCATACGCTAGCTTTTGTGTAATTTCTTGGACCCAGAAATTTGTCCGCCGA 2758
Qy |||||
Db 3363 CGAGGAATTCACAAAGGACAAAGGTTTGATT 3393
Qy |||||
Db 2759 CGAGGAATTCACAAAGGACAAAGGTTTGAAT 2789
Qy |||||

RESULT 7

US-10-209-059-25
; Sequence 25, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3470
; TYPE: DNA
; ORGANISM: Zea mays
; US-10-209-059-25

Query Match 45.4%; Score 1563.4; DB 16; Length 3470;
Best Local Similarity 70.9%; Pred. No. 0;
Matches 2369; Conservative 0; Mismatches 671; Indels 303; Gaps 10;

Qy 97 GGCACGCGCGCGCTGGTGGCCCGCTCGCAACCCGGAACGAGCTGTGTGATCCGGGG 156
Db 43 GGTACCGGTGGCTCGCCCGGCTCGCACATGCGGAGAGCTGATGTCATGTCGCGC 102
Qy 157 CCACGAGGACCCCAAGCC---GCTGGGGCGCTGAGCGGGCAGGTGCGAGATATCGG 213
Db 103 CCAGGAGGAGCCAAACGCAAGGTCCGAGCGCGAGCTGAAGACGTCGCGTGGCGC 162
Qy 214 CGACGAGGTCCGCTCAACGCTGCAACGCGACCTTTCGTCGCTGCAACGAGTCCGCTT 273
Db 163 CGACGAGGTCCGAGCGGAGGAGCGGCGAGCCCTTCGTCGCTGCGCGAGTGGCGCTT 222
Qy 274 CCCCGTGTCCGCCCTTGTCTACAGTACGAGCGCCGGGAGGCAACGAGAACTGCCCCCA 333
Db 223 CCCCGTGTCCGCCCTTGTCTACAGTACGAGCGCAGCGAGGCAACGAGTGGCGCGCA 282
Qy 334 GTCGAGACGCGCTACAGCGCTCAAGGGAGCCCGAGGGTTCGCGGAGCGATGACGA 393
Db 283 GTCGAGACCGCTACAGCGCCAGAAAGGTTGCCGAGGTTGCCGAGGAGGAGGA 342
Qy 394 GG---AGGACATCGACGACCTGAGCAGCAGTTCACATCGACGAGAAATCAGCAGAG 450
Db 343 GGGCCCGGAGATGAGACGACTTCGAGGACGAGTTCCCGCCCAAGAGCCCTCA 402

QY 451 GCAGCTGGAGGCAACATGACAGAACAGCCAGATCACAGAGCGCATGCTGCACCGCAGGAT 510
Db 403 CGAGCTGTGCGCTTTCGACCTCTACTCGAGAAACGCGCAGCAACCGCGCGCAGAAATGGCG 462
QY 511 GAGCTAGGAGGGGCCCGACGACGGCGACGCAACACCCCGCAGATCCGCGCCAT 570
Db 463 GACGGTGG-----CCAGACGCTGTGCTC 486
QY 571 CATCACCGGCTCCGCTCCGTCGCGGTGAGCGGTGAGTTTCGATTACCAACGGGTATGG 630
Db 487 CTTTACCGGAAGCTCGCCGGGAAG----- 512
QY 631 CCAGGGGAGGTCTGCTTTCCTTCGACAAAGCGATCCATCCGTATCCCTGTGTCTGAGCC 690
Db 513 -----ACCTGAGGCGGAGAG 528
QY 691 AGGAGTGCACAGTGGGACGAGAAGAAAGTGAAGTGAAGGAGGATGACCACTG 750
Db 529 GGAGATGGAGGAGCAATGGAGTGGAGGACCGGATCGACAAGTGAAGACCAAGCAGGA 588
QY 751 GAAGTCCAAAGCAGGGCATCTCGCGCGCGCGCGCATCCCGAAGACATGGACCGCGACGT 810
Db 589 GAGAGGGCAAGCTCAACCAACGACGACGACGACGACGACGACGACGACGACGACGACG 648
QY 811 G----GCACCTGAACGACGAGCGAGGACGCGCTGTGCGAGGAAGGTGTGATTCGCTCGAG 867
Db 649 GTACATGCTGCTTTCGCGAGCGCGACAGCGCGTGTGCGCAAGGTTCCGATCCCGTCGAG 708
QY 868 CAAAGTGAAACCCGTACCGGATGTGATCGTGGTGGCTCTCGTTGTCTCGCTTCTTCTCT 927
Db 709 CATGATCAACCCGTACCGCATGCTCATGCTGCTCCGCTGCTGCTGCTGCTGCTTCTTCT 768
QY 928 CCGGTACCGTATCTGCAACCCGCTCCCGAGCGCATCGGGCTGTGGCTGCTCTCCATCAT 987
Db 769 CAAGTTCGGATACAGACGCCCGCACGAGCGCGTGCCTCTGTGGCTGGCTCCGTCAT 828
QY 988 CTCGGAGATCTGTTGCGCATCTCTGATCTCTGATCTCTGACCAAGTTCGCCCAT 1047
Db 829 CTCGGAGCTGTGTTGCGCTTCTCTGATCTCTGATCTCTGACCAAGTGGCGCGCT 888
QY 1048 CGACCGGAGACCTACTCGACCGCTCTCCCTCAGGTACGAGAGGAGGGAGCGCTC 1107
Db 889 GACCGGGAGACGTACTGACCGCTTGGCGCTGCGGTACGACCGTGAAGGCGAGCGGT 948
QY 1108 GCTGCTGTGCGCGGTGAGCTGTTCGTGAGCAGGTGACCCCGCTCAAGGAGCGCGCT 1167
Db 949 CCGGCTGTCCCATCGACTTCTTGTGACAGGTGACCCCGCTCAAGGAGCGCGCAT 1008
QY 1168 GGTGACCGCAACACGCTGCTCTCATCTCGCGGTAGACTACCCCGTGAACAGGTCTC 1227
Db 1009 CATCACCGCAACACCGTGTGTCTCTCTCGCGCTCGACTACCCCGTGAACCGCTCAG 1068
QY 1228 CTCGTAGCTCTCGACGAGCGCGCTCGATGCTGACGTTCGCTGCTGCGAGACGCG 1287
Db 1069 CTCGTAGCTCTCGACGAGCGCGCTCATGCTGCTCTCGACGCTGTTCGAGACCGC 1128
QY 1288 CGAGTTTCGCGCAAGTGGTGCCTTCTGCAAGAAAGTTCGGCATCGAGCCCGCGCCCC 1347
Db 1129 CGAGTTTCGCGCGCGCTGGTGCCTTCTGCAAGAAAGTTCGGCATCGAGCGCGCGCCCC 1188
QY 1348 GGAAGTCTACTTCTCGTCAAGTTCGACTACTCAAGGACAAAGTTCAGGCCCACTTCGT 1407
Db 1189 GGAAGTCTACTTCTCGCAGAAAGTTCGACTACTCAAGGACAAAGTTCAGGCCCACTTCGT 1248
QY 1408 GCAGGAGCGCGCGCATGAGAGAGATGATGAGGATTCAGGTCCGATCAACGGCT 1467
Db 1249 CAAGGAGCGCGCATGAGAGAGAGATGAGGATTCAGGTCCGATCAACGGCT 1308
QY 1468 GGTGGCCAAAGCCATGAAGTTCGCGCAGAGGGGTGGATCATGAAGGACGGCACGCGCTG 1527
Db 1309 GGTGGCCAAAGCGCAGAGAGAGCCGAGGAGGGTGGTCAATGACGAGCGCACGCGTG 1368
QY 1528 GCCCGGGAACAACACCGCGACCAACCCCGCATGATCCAGGTGTTCTCGGGCCACAGCGG 1587

Db 1369 GCCCGGGAACAACACGCGGACACCCCGGTATGATCCAGGTCTACTCTCGCAACCCAGGG 1428
QY 1588 CGGCGCACGACACCGAGGCAACGAGCTGCCCGCTCGGTAGCTCTCCCGTGAGAGCG 1647
Db 1429 CGGCTGGAGCTGGAGGGCCACGAGCTGCCCGCTCTGTCTAGCTCTCCCGTGAAGCG 1488
QY 1648 CCGGGATTCCAGACACCAAGAGGCGCGGCCATGAACGCTCTGATTCGGGTCTCCCG 1707
Db 1489 CCGCGGTACAACCAACCAAGAGGCGGCGCCATGAACGCGCTGGTGGGTCTCCCG 1548
QY 1708 CGTGTGACCAACGCGCCATTCATGCTCAACTTTGATCTGTGATCTACTATCAACACAG 1767
Db 1549 CGTGTCTACCAACGCGCCCTTCATCTCAACCTCGACTCGACCACTAGCTCAACACAG 1608
QY 1768 CAAAGCCATCCGGGAGGCGATGCTCTCATGACCTCAGGTGCGCGGAAGCTCTG 1827
Db 1609 CAAAGCGGTGCGGAGGCGATGCTCTCTCATGACCCGAGCTGGGGAAGAGCTCTG 1668
QY 1828 CTACGTTCAAGTTCGCGCAGAGGTTCCAGCGCATCGACGTGACGACCGGATACGCTAACAG 1887
Db 1669 CTACGTTCAAGTTCGCGCAGCGCTTCGATGCGATCGATCGCCACGACCGATACGCCAACCG 1728
QY 1888 GAAACCGCTCTTCTTTCGACATCAACATGAAGGGCTGGAACGATCCAAAGGCCGCTGTA 1947
Db 1729 CAAAGCTGCTTCTTTCGACATCAACATGAAGGGCTGGAACGATCCAGGGCCGCTGTA 1788
QY 1948 CGTGCGGACAGGCTGCTTCGCGCGCAGCGCTCTACGGCTACAACTCCCAAGGG 2007
Db 1789 CGTGCGGACAGGGGTGCTGTTCAACCGCCAGGCGCTGTACGGCTACGACCCGCGCGGCC 1848
QY 2008 ACCCAAGAGGCCCAAGATGCTGAC-----TGCGACTGCTGCCCGCTG 2049
Db 1849 CGAGAAGCGGCCCAAGATGACGTGCGACTGCTGCGCGTGTGCTGCTGCTGCTGCTG 1908
QY 2050 CTTTGGCGCCAAAGAGCG----- 2067
Db 1909 CTTTGGCGCGCAAGCGCGCAAGCGCGCAAGGCGCAAGAACAGAGGCGCGCGCGAGGA 1968
QY 2068 -----GAAACACGCCAAGACCG 2085
Db 1969 GCGCGCGCGGGCTGCTCGGCTTCTACAGGAAGCGGACCAAGAGGAACAGCTCGCGG 2028
QY 2086 GCTGCGGAGGCGCACCGCTGATATGGAGTAGATAGCAAGAG----- 2131
Db 2029 CGGTCGTTGGCGGCGAGCAAGAGGCGCGGCTGTACAAAGAACACACAGCGCGCTT 2088
QY 2132 -----ATGCT 2136
Db 2089 CGAGCTGGAGGAGATCGAGGAGGGCTGAGGGGTACGACGAGCTGGAGCGCTCTCGCT 2148
QY 2137 CATGTCCACATGAACCTTCGAGAGCGGTTTCGGGAGTCCGCGGCTTCGTCAAGCTCGAC 2196
Db 2149 CATGTGCGAAGAGCTTCGAGAGCGGTTTCGGCCAGTCCGCGGTTCATCGCTCCAC 2208
QY 2197 GCTGATGAGGAGGCGCGCTCCCTC-----CTTCTGTCGAGCCCGCGCTCCTCAA 2250
Db 2209 GCTGTCGAGGAGCGGGCTTCGCGAGGCGCGCGCGCGCGCGCTCATCAA 2268
QY 2251 GGAGGCGCATTCATGCTCAGCTGCGGCTACGAGGACAAAGACGACTGGGGGGTGGAGCT 2310
Db 2269 GGAGGCGCATTCAGCTCATCAGCTGCGGATACGAGGAGAAAGCCGAGTGGGCGAAGGAGAT 2328
QY 2311 GGGGTGGATCTACGGGTTCATCGAGGACATCTTGAAGGGGTTCAAGATGACATGCGCG 2370
Db 2329 TGGGTGGATCTATGGGTTCGTCAGAGGATATCTTGAAGGGGTTCAAGATGACATGCGCG 2388
QY 2371 CCGGTGGCGCTCCGCTGATCTGATCCGAGAGGGCGCGGTTCAAGGGGTTCGCGCGCGAT 2430
Db 2389 GGGGTGGAAGTCCGCTGATCTGCAACCGCACGCGCGCGGTTCAAGGGGTTCGCGCGCGAT 2448
QY 2431 CAATCTATCGGACCGTCTCAACAGGTGCTCCGTTGGGCGCTGGGGTTCGTGAGATCTT 2490

Db 1559 GCCAGGGAATAACAATAAGGATCAATCTGTGTATGATTAACAAGTGTTCCTTGGTAGCAGTGG 1618
Qy 1588 CGGCCACGACACGAGGCGCAACGAGCTGCCCGCTCGTGTACGTCTCCCGTGGAGAGCG 1647
Db 1619 AGGTCTTGATCTGAAGGAACCAACTCTCTCGCTGTGTATGTTTCCAGAGAGAAAG 1678
Qy 1648 CCGGGGATTCAGACACCAAGAAGGCGCGCCATGAACGCTCTGATTCGCGTCTCCG 1707
Db 1679 GCCTGGTCTTCAACACCAAGAAGGCTGGTGCCATGAATGCTCTGGGTATCTGC 1738
Qy 1708 CGTCTGACCAACGCGGCATTCATGCTCAACTGGAGTGTGATCACTACATCAACACAG 1767
Db 1739 TGTCTCACAAATGCTCTTTCATGTTGAATCTGGATTTGATCACTATGTCATAACAG 1798
Qy 1768 CAAGGCCATCCGGGAGGCCATGTGCTTCTCATGGAACCTCAGGTCCGCGGAAGTCTG 1827
Db 1799 CAGGCTGCCGAGAGGCCATGTGCTTCTTGATGGACCCCAAACTGGGAAGAGTCTG 1858
Qy 1828 CTAGTTCAGTTCGCCGAGAGGTTTCGACGCGATCGAGTGCACGACCGATACGCTAACAG 1887
Db 1859 CTATGTCAGTTCCTCAAAGATTTGATGATGATTAATGATACACATGATCGTTATGCCAACAG 1918
Qy 1888 GAACACGCTCTCTTCGACATCAACATGAAGGGCTGGAGGCGATCCAAGGCCGGTGA 1947
Db 1919 GAACACAGTCTTCTTTCATATTAACATGAAGGGCTAGATGGTATTCAGAGTCTCTGTATA 1978
Qy 1948 CGTCCGACAGGGTGTGCTGTTCCGGCGCCAGGCGCTCTACGGCTACAAACCTCCCAAGG 2007
Db 1979 TGTGGGACTGGATGTGTTTCAGGAGCGAAGCTTTGTATGGCTATTAATCTCCCAAGG 2038
Qy 2008 ACCCAAGAGCCCAAGATGTGACCTCGACTGCTGCCCGTGTTCG ---GCCCGAAGAA 2064
Db 2039 TCCAAAGCGTCCAAATAATGTAGCTGTGATTTGTGCCCGTGTTCGGAAGCGCAAGAA 2098
Qy 2065 GCGGAACACGCCAAGAGCGGGTCCGAGGGGCAACGCTGATATGGAGTAGATAGCGA 2124
Db 2099 GTATAAGGAAGAATGATGCAATGGAGAGGCTGCAAGGCTAAAGGGATGATGATGA 2158
Qy 2125 CAAGGAGATGCTCATGTCCACATGAATCTCGAGAAGCGGTTCCGGCAGTCCCGCGGCT 2184
Db 2159 CAAAGAGTGTGATGTCCTCAATGAATTTGAGAGAAATTTGGACAACTCTATTTT 2218
Qy 2185 CGTCACGTGACGCTGATGAGGAAGCGCGCTCCCTCTCTGTCGAGCGCCCGCGCGCT 2244
Db 2219 TGTGACTTCTACCTTGTATGGAAGAGGGTGTGTGCTCTCTTCTCAAGTCCAGCTGCCCT 2278
Qy 2245 CCTCAAGGAGGCCATCCATGCTCAGCTGGGCTAGAGACCAAGACCGACTGGGGCT 2304
Db 2279 GCTTAAAGAAGCCCAATCATGTGATTAGCTGTGGATATGAAGATAAAACTGAATGGGGACT 2338
Qy 2305 GGAGCTGGGTGGATCTACGGGTCCGATCAGGAGGACATCTGACGGGGTTCAAGATGCA 2364
Db 2339 TGAGCTGGTTGGATCTATGATCTATCAAGAGATATTTCTAACAGTCTTTAAGATGCA 2398
Qy 2365 CTGCGCGGGTGGCGCTCCGCTGATCGATGCCGAAGCGCGCGCTTCAAGGGGTCCG 2424
Db 2399 TTGCGGTGGGTGGAGTCCATTTATGTATGTCGAAGAGAGCTGCAATCAAGGGTACTGC 2458
Qy 2425 GCCGATCAATCTATCGACCGCTCTCAACAGGTGCTCGGTGGCGCTGGGTCCGTCGA 2484
Db 2459 TCCTATCAACTGTTCAGATCGTCTCAACCAAGGTTCTTCGTGGGCACCTTGGTTCATTTGA 2518
Qy 2485 GATCTTCTTCAGCGCGGACACGCCCTCGCTGTACGGCTACAAGAACGGCAACCTCAAGTG 2544
Db 2519 GATTTCTTTAGTACCACTTGGCTCTATGTTGGCTTCAAGGAAGAAGCTAAGTG 2578
Qy 2545 GCTGGAGCGTTCGCCCTACATCAACACCAATACCCCTTCACTCGCTCCGCGTCT 2604
Db 2579 GCTTGAGAGATTTGCCCTATGCAAAACAACACTGTCTATCCATTCACCTCCATTCCTCTAGT 2638
Qy 2605 CGCTACTGACCTCCCGCGCTCTGCTCTCCACCGGCAAGTTCAATGCGCGTAT 2664
Db 2639 TGCCTACTGTATCTTCCAGAGGTTTGTTTACTCACTGACAAATTCATCATGCCACCGAT 2698

Qy 2665 TAGCAGTTTCGCGAGCCTCTTCTTTCATCGCCCTCTTTCATGTCCATCTTCGCGACGGGCAT 2724
Db 2699 TAGCAGCTTTGCTGGTTTGTACTTTGTCTCTTCTCTCAATCATTTGCAACTGGTAT 2758
Qy 2725 CTTGGAGATCGCGTGGAGCGGGTGAAGATCGAGAGTGGTGGAGAAAGAGAGTTCGT 2784
Db 2759 TCTTGAGTTCAAAATGGATGGAGTGAAGTGAAGAAATGGTGAGAAATGAGCAGTTTTG 2818
Qy 2785 GGTTCATCGCGGGGTCTCGGCATCTCTTTCGCGCTGTCGAGGGCTGCTCAAGGTCT 2844
Db 2819 GGTTCATTTGGTGGTATCAGCTCACCTCTTGTCTGTATACAGGTCTGCTAAAGGTCT 2878
Qy 2845 CGCGGGATCGACCAACACTTTCACCGCTCACCTCCAAAGGCCACCGCGGAGAGGACGACGA 2904
Db 2879 GGCTGGAATTGACCACTTTCACCTGTATACATCAAGGCAAC-----AGATGATGAAGA 2932
Qy 2905 GTTTCGCGAGCTCTACGCTTCAAGTGGAGACGCTCTCATCTCCGCGCCACGAGTCT 2964
Db 2933 GTTTGGAGAAATGTACACCTTTAAGTGGACTACACTCTTTGATTTCTTCCCAACCACTATTT 2992
Qy 2965 CATCATTAACGTCATCGGCTGTGGCGGCGCATCTCCGACGCCATCAACAACGGGTACCA 3024
Db 2993 GATCATTAACATTTGTTGGTGTGTTGCTGGAATCTCAGATGCCATAAACAATGGGTACCA 3052
Qy 3025 GTCTCGGGGGCCCTCTTCGGCAAGCTCTTTCGCTTCTTTCGCTTCTTGGGTCTATGTCACCTCTA 3084
Db 3053 ATCTCGGGGACCACTCTTTGGAAAGCTCTTCTTCTTCTTGGTGTATGTCCTCTCTA 3112
Qy 3085 CCGCTTCTCAAGGGGCTCATGGGGGCCAGAACAGGACGCCACCGTGTGTTGTCTATCTG 3144
Db 3113 TCCATTTCTTAAAGGTTTGAATGGGTGCCAAATCGCAAAATCGCACACCACTTGTGTGTTG 3172
Qy 3145 GTTCATTTCTGCTGGCTCCATCTTCTCCCTGCTCTGGGTGAGGATCGACCTTTTCATCGT 3204
Db 3173 GTCAGTGTCTATTGGCTCTTATTTCTCTTCTTCTTCTTGGTAAAGATTTGATTCATTTGCTCT 3232
Qy 3205 CAGGACCAAGGGGCCGAGCGTTCAGGCGAGTGTGGCATCAATTTGCTGA 3250
Db 3233 CAAGACTAAGGGACCTGATACCAAGCTATGTGGAATCAACTGCTAA 3278

RESULT 10
US-10-437-963-83811/c
; Sequence 83811, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83811
; LENGTH: 3493
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRF4530_83107C.1
US-10-437-963-83811

Query Match 41.8%; Score 1439.6; DB 18; Length 3493;
Best Local Similarity 69.4%; Pred. No. 0;
Matches 2296; Conservative 0; Mismatches 669; Indels 345; Gaps 11;

QY 187 GAGCGGCGAGTGTGCGGATATGCGGCGACGAGGTGCGGCTCAGGTGGAGCGGACCT 246
Db 3455 GAAAGAGAAGACGTGCCGGGTGTGCGGCGAGAGGTGCGGAGGGGAGACCGGAAGCC 3396
QY 247 CTTTCGTCCTGCAACGAGTGGCTTCCCGTGTGCGGCGCTGCTACGAGTACGAGCG 306
Db 3395 GTTCGTGCGGTGCGGCGAGTGGGCTTCCCGTGTGCAAGCCCTGCTACGAGTACGAGCG 3336
QY 307 CCGGAGGGGCGACGAGAACTGCCCCAGTGAAGACGCGCTACAAAGCGCTTCAAGGGGAG 366
Db 3335 CAGCGAGGGGCAACCAAGTCTGCCCCAGTGAACACCCCGCTACAAAGCGCCCAAAAGGGTG 3276
QY 367 CCCGAGGGTTCGCGGAGCGATGCGA---GGAGGACATCGACGCTGGAGCGAGT 423
Db 3275 CCCACGGGTGAAGCGGACGAGGACGACGCGGCGACATGACGACTTCGAGGGAGGATT 3216
QY 424 CAACATCGACGACGAGAAATGACGAGGCGAGCTGGAGGGGCAACATGACGAAACAGCCAGAT 483
Db 3215 CCAGATCAAGAGCCCCACCAAGCAGA-----AACCCCCCACGAGCCCGTC 3170
QY 484 CACGAGCGGATGTGCGACGGGAGGATGAGCTACGGGAGGGGCCCCGACGACGCGGACGG 543
Db 3169 AACTTCGACGCTACTCTCGGAGAAACGGCGAGCAGCGGACAGAAAGTGGCGCCTGGAGG- 3111
QY 544 CAACACACCCCGCAGATCCCGCCCATCATACCGGCTCCCGCTCGGTGCGGCTGAGCGG 603
Db 3110 -----CCCGGCGCTCTCTTCCCTTCAACCGGAAGCGTGGCTGGGAGG- 3070
QY 604 TGAGTTTCCGATTACCAACGGGTATGGCCACGCGGAGGTCTCGTCTTCCCTGCAACAGCG 663
Db 3069 ----- 3070
QY 664 CATCCATCCGTAACCTGTGTCTGAGCCAGGAGTGCCTAAGTGGGACGAGAAAGAAAGT 723
Db 3069 -----ATCTGAGCAGGAGAGGAGATGGAGGGTGGCAT---GGAGTGAAGGACAGGAT 3018
QY 724 GAGCTGGAAGAGAGGATGAGGACGACTGGAAGTCCNAGCAGGGCATCTCCGCGCGGCGGC 783
Db 3017 CGACAAGTGGGAAGCAAGAGCAGGAGAAAGCGGGGCAAGCTCAACCGCGACGACAGCGACA 2958
QY 784 CGATCCCGAAGACATGACGCGGAGCTGGCACTGAAACGACGAGCGGAGCGCGCTGTC 843
Db 2957 CGACGACGACAAAGAACGACGAGTACATGCTGCTCGGAGGCGGAGGCGGCTGTG 2898
QY 844 GAGAAAGTGTGATTCGCGTCCGAGCAAGGTGAACCCGTAACCGGATGATGATGCTGTCG 903
Db 2897 GAGAAAGTGTCCGATCCCGTCCGAGCAAGATCAACCCGTAACCGGATGATGATGCTGTCG 2838
QY 904 TCTGTTGTGCTCGGCTTCTTCTCCGTAACGCTATCTGACACCCGTCGCGGAGCGCAT 963
Db 2837 GCTGGTGTGCTGCTGCTTCTTCTCAAGTTCCGGATCAACGCGCGCGATGGAACGCGGT 2778
QY 964 CGGCTGTGGTCTGCTCCATCATCTGCGAGATCTGTTTCCGCTATCTCTGGATCCTCGA 1023
Db 2777 GCCGCTGTGGTCTGCTGATCTGCGAGCTGTGTTTCCGCTGTGCTGTGATCTCTCGA 2718
QY 1024 CCAGTTCCTCAAGTGGTTCCTCATGACCGGAGAGCTGACCTCGACGCGCTCTCCCTCAG 1083
Db 2717 CCAGTCCCAAGTGGTTCGCGGTGACGAGGGGAGAGCTACCTGAGCGGCTGGCCCTCG 2658
QY 1084 GTACGAGGGGAAGGGGAGCGTCTGCTGTGCGCGGTGACACCTGTTCTGAGGACAGGT 1143
Db 2657 GTACGAGCGGACGCGGAGCGGTGCGGCTGCGGCTGCGGCTGCGATCTTCTGTCAGCAGGT 2598
QY 1144 GGACCCGCTCAAGGAGCGCGCTGCTGACCGCCCAACACCGTGTCTTCCATCCTCGCGGT 1203
Db 2597 GGACCCGCTCAAGGAGCGCGGCTGCTGACCGCCCAACACCGTGTCTTCCATCCTCGCGGT 2538
QY 1204 AGACTACCCGCTGGAACAAGTCTCCTGTGATGCTTCCGACGACGCGCGCTGAGTCTGAC 1263
Db 2537 CGACTACCCGCTGACCGCGCTCTCTGTGATGCTTCCGACGACGCGGCGTCTCATGCTGCT 2478
QY 1264 GTTCGAGTGTGTCGAGAGCGCGGAGTTCGCGGCAAGTGGGGTGGCCCTTCTGCAAGAA 1323

Db 2477 CTTTGCACACGCTCTCCGAGACCGCCGAGTTTCGCCCGCGGTGGGTCCCTTCTGCAAGAA 2418
QY 1324 GTTCGGCATCGAGCCCGCGCGCCCGGAGTTCTACTTCTCGCTCAAGGTTCGACTACCTCAA 1383
Db 2417 GTTCACCATCGAGCCCGCGCGCCCGGAGTTCTACTTCTCCAGAAGATCGACTACCTCAA 2358
QY 1384 GGACAGGTGAGCCCACTTCTGTCAGGAGCGCGGCCCATGAAAGAGAGATGAGGA 1443
Db 2357 GGACAGGTGAGCCCACTTCTGTCAAAGAAACGCGCGGCCATGAAAGAGAGATGAGGA 2298
QY 1444 GTTCAAGGTCCGAGTCAACCGCTGTGGCCAGGCGCATGAAGTGCAGGAGGGGTG 1503
Db 2297 GTTCAAGGTGAGGATAAACCGCTGTGGTGGGAGCGCGAGAGAAAGCCGAGGAGGGTG 2238
QY 1504 GATCATGAAGGACGCGCACGCGTGTGGCCCGGAAACAACCCGCGACCAACCCCGCATGAT 1563
Db 2237 GGTGATGACGAGCGGACGCGCATGCGCGGGAAACAACAGGAGCACACCCGGGATGAT 2178
QY 1564 CCAGGTGTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCAACGAGCTGCCCGCT 1623
Db 2177 CCAGGTGATCCTTGGGACAGCCAGGCGCGCTCGACGTCGAGGGCAGCGCTGCCGCGCT 2118
QY 1624 CGTGTAGTCTCTCCGTCGAGAGCGCCGCGGATTCACGACACCAAGAGGCGCGGCCAT 1683
Db 2117 GGTGTAGTGTCTCCGCGAGAGGCGCCCGGCTACAACCAACAGAGGCGCGGCCAT 2058
QY 1684 GAAACGCTCTGATTCGCGCTCTCCGCTGTGCTGACCAAGCGGCATTCATGCTCAACTTGA 1743
Db 2057 GAACCTCCTGTTGCGCTCTCCGCGGTGCTTACCAACGCCCTTTCATCTCAACCTCGA 1998
QY 1744 CTGTGATCATATACATAACAACAGAGGCGCATCCGGAGGCCATGCTTCTTCTCATGGA 1803
Db 1997 CTGCGACCATCTACGTCACCAACAGAGGCGCTCCGGAGGCCATGCTTCTTCTCATGGA 1938
QY 1804 CCCTCAGTGTGGCGGAGGTCTGCTAGTTCAGTTCGCGCAGAGGTTCCAGCGCATCGA 1863
Db 1937 CAAGCAGCTCGGCAAGAGCTGTGCTAGTTCAGTTCAGTTCAGCGGCTTCGACGCGCATCGA 1878
QY 1864 CGTGCACGCGCATACGCTAAACAGAAACACCGTCTTCTTTC----- 1903
Db 1877 CCGCCACGATCGCTACGCCAACCGCAACCGCTCTTCTTCGACGTGAGCATGCGCTTCTT 1818
QY 1904 -----GACATCAACATGAAGGGGCTGACGCGCATCCAAAG 1938
Db 1817 CTCGCGCTGCTGCTGCTGCGCGCTGACATCAACATGAAGGGGCTGACGCGGATACAGG 1758
QY 1939 CCGGCTGTGCTGCGGACAGGTCGCTGTTCCGCGCGCAGCGCTCTACGCTACACCC 1998
Db 1757 GCGGTTGTGCTGCGGACGCGGACGCTGTTCAACAGGACGCGCTGTACGATACGACCC 1698
QY 1999 TCCCAAGGGACCCCAAGAGGCCCAAGATGCTGAC-----TGCGACTG 2040
Db 1697 GCGCGCGCGGAGAGAGGCGGAGATGACGTGCGATGCTGCGCGCTCGTGGTGTGCTG 1638
QY 2041 CTGCGCGCTGCTGCGCGCGCAAGAGCG----- 2067
Db 1637 CTGCTGCTGCTTTCGCGCGGCGGAGCGCGCAAGTTCGCAAAAGAGGGCGCGG 1578
QY 2068 ----- 2067
Db 1577 CCGCGCGGAGGCGCGGCTTCGACGAGCGCGCGCGGCTGCTCGGGTTCACAGAA 1518
QY 2068 -----GAAACAACGCAAGGACGCGCTCCGAGAGGCGCACCGCTGATATGGGAGTAGATAGC 2122
Db 1517 GAGGAGGCAAGAGGACAGCTCGCGCGCGGCGCGCTGCTGCTCGCGGAGGAGAAAGG 1458
QY 2123 GACAAGGA----- 2130
Db 1457 GTACCGGAAGCACCGAGCGCGGTTTCGAGCTGGAGGAGATCGAGGAGGCGCTCGAGGGGTA 1398
QY 2131 -----GATGCTCATGTCCCATGAACTTCGAGAGCGGTTTCGGGCA 2172

Db 1397 CGACGAGCTGAGGGCTCGTGGCTCATGTGCGAGAGAGCTTCGAGAGCGGTTGGGCCA 1338
Qy 2173 GTCCGGCGGTTGTCATCGCTCGACGCTGATGGAGAGAGGGCGTCCCTC-----CTTC 2226
Db 1337 GTGCGGGTGTTCATCGCTCCACCTCGTCGAGGACGGCGCTCCCCAGGGCGCGC 1278
Qy 2227 GTGAGCCCCCGCGCTCTCTCAAGGAGGCGATCATGTATCATAGCTGCGGCTACGAGGA 2286
Db 1277 CGCGACCCCGCGCTCTCATCAAGGAGGCGATCATAGCTGCGGCTACGAGGA 1218
Qy 2287 CAAGACCGACTGGGGCTGAGCTGGGCTGATCTACGGTTCGATCATCGGAGGACATCCT 2346
Db 1217 GAAGACCGAGTGGGGCAAGAGATTGGTGGATCTACGGTTCGATCGGAGGACATCTT 1158
Qy 2347 GACGGGTTCAAGATGACCTCGCGGGTGGCGTCTCGTGTACTGATGCGGAGCGGCG 2406
Db 1157 AACGGGTTCAAGATGATTGGCGTGGTGGAGTTCGTGTACTGACGCGCGGAGGC 1098
Qy 2407 GGGTTCAAGGGTTCGGCGCGATCAATCTATCGGACCGTCTCAACAGGTCTCGGGT 2466
Db 1097 GGCATTCAAGGGTTCGGCGCGCATCAACCTGTGGATCGTCTGCACGAGTGTCTCGGT 1038
Qy 2467 GGGCTGGGTCGGTTCGAGATCTTCTTCAGCGGACAGCCCCCTGTGTACGGTACAA 2526
Db 1037 GGGCTGGGTCGGTTCGAGATCTTCTTCAGGCGCGCATTTGCCGCTCTGTACGCT---A 981
Qy 2527 GAACGGCAACCTCAAGTGGTTCGAGCGCTTCGCCATACATCAACACCACTCTACCCCT 2586
Db 980 TGGCGCGCTCAAGTGGTTCGAGCGCTTCGCCATACCAACCACTCTCTACCCCT 921
Qy 2587 CACTCGCTCCGGTTCGCTACTGCAACCTTCGCCGCTTCGCTCTCAACCGGCAA 2646
Db 920 CACTCCTCACTCCCTTCGCTACTGCAACCTTCGCCGCTTCGCTCTCAACCGGCAA 861
Qy 2647 GTTCATCATCGCTTCGATAGCGTTCCGCGCTCTTCTTCATCGCCCTCTTCATGTC 2706
Db 860 GTTCATCATCCCAACGCTTAACTTTGGCGAGATATGTTTATAGCGTTTTCCTGTC 801
Qy 2707 CATCTTCGACCGGCACTCTCGAGATGCGGTGGAGCGGCTGAGCATCGAGAGTGGTG 2766
Db 800 GATCATCGACGGGGTGTGGAGCTGCGTGGAGCGGGTGGAGCATCGAGACTGGTG 741
Qy 2767 GAGGAACGAGAGTTCTGGGTTCATCGCGGGTGTTCGGCGATCTCTTCGCGCTGTCGA 2826
Db 740 GAGGAACGAGAGTTCTGGGTTCATCGCGGGTGTTCGGCGCATCTGTTCCGCTGTTCCA 681
Qy 2827 GGGCTGCTCAAGTCTCCCGGATGACACCACTTCACCGTCACTCCAGGCCAC 2886
Db 680 GGGCTGCTCAAGTGTCTGGCGGCTGGACCACTTCACGTTGACGTTCGAAGGCCG 621
Qy 2887 CGGCGAGGAGACGACGAGTTTCGCGAGCTCTACGCTTCAAGTGGACCACTCTCTCAT 2946
Db 620 CGCGAGAGACGACGCTTCGGGAGCTCTACCTGTTCAAGTGGAGAGCTGTGGT 561
Qy 2947 CCGCCCAACGCTGCTCATATTAACGTATCGGGTGTGGCGGATCTTCGAGCG 3006
Db 560 CGCGCCACGACGCTGATCATCAACATGTTGGGATGCTCGCCGCGTTCGAGCG 501
Qy 3007 CATCAACAGGGTACAGTCTTCGGGGCCCTCTTCGCAAGCTCTTTCGCTTCG 3066
Db 500 CGTGAAACAAGGGTACGGGTTCGGGGCCCGCTGTTTCGGAAGCTCTTCTCTCTG 441
Qy 3067 GGTTCATCGTCCACTTACCGTTTCCTCAAGGGGCTCATGGGGCGCCAGAACAGGAGCC 3126
Db 440 GGTTCATCTCACTTACCTTACCTTCCTCAAGGGGCTCATGGGAGGAGAACCGAGGCC 381
Qy 3127 CACGGTTGTTGTCATCTGGTTCATTTGCTGGGCTTCATCTTCTCTGCTTGGGTGAG 3186
Db 380 CACCATTTGCTGCTGTCCTTCCTTCCTCGCTTCCTCTCTCTCTCTCTCTCTCTCT 321
Qy 3187 GATCGACCTTTCATCTGAGGACCAAGGCGCGGAGTTCAGGAGTTCGATCAATTG 3246
Db 320 GATCGACCTTTCATCTCCCAAGCCCAAGGGCCCGTCTCTCAAGCCATGCGGGGTCTCGT 261

Qy 3247 CTGAGCTGTT 3256
Db 260 CTGAGCTGCT 251
RESULT 11
US-10-160-719-29
; Sequence 29, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3551)
US-10-160-719-29

Query Match 40.8%; Score 1397.6; DB 16; Length 3746;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;
Qy 133 GAACGAGCTGTGCTCATCCGGGGCCACGAGGACCCCAAGCGCTCGGGCGCTGAGCGG 192
Db 308 GAGCTCGTTGCCATGGAGGGCGACCGGACGCGGTGAGTCTGGGAGGGCGCGTGGCGG 367
Qy 193 GCAGGTGTGCGAGATATGCGGCGACGAGGTTCGGGCTCACGGTGGACGCGACCTTTCTG 252
Db 368 ACAGGTGTGCGAGATCTCGGCGACGCGTGGGACACACGCGGAGGGGACGCTTTCG 427
Qy 253 CGCTTCGAACGAGTGGGCTTCCCGTGTGCGGGCCCTGCTACGAGTACGAGCGCGGGA 312
Db 428 CGCTTCGACGCTGCGGGGTTTCGGTGTGCGGGCCCTGCTACGAGTACGAGCGCAAGGA 487
Qy 313 GGGCACGAGACTGCCCCAGTGAAGACGCGCTACAAGCGCTCAAGGGGAGCCGAG 372
Db 488 CGGCAACGAGCGTCCCCAGTGAAGACCAAGTGAAGCGCCACAGGGGAGCCCGC 547
Qy 373 GGTTCGCGGGAGCATGACGAGGAGGACATCGACGACCTGGAGCACGAGTTCAACATCGA 432
Db 548 GATCCGTGGGAGGAGGAGACGACACTGATGCCGATAGCGACTTCAATTACCTTGCATC 607
Qy 433 CGACGAGATACGACGAGGAGCTGGAGGGCAATGCGAGAACAGCAGCATCA----- 485
Db 608 TGGCAATGAGGACACAGAAGCAGAAGATTGCCGACAGATGCCGAGTGGCGCATGAACGT 667
Qy 486 -----CCGAGGGGATGTGCGCGGAGGATGAGTACGGGAGGGGCCCG----- 530
Db 668 TGGGGGAGCGGGGATGTTGGTCCGCCCAAGTATGACATGAGTGGCGAGATCGGGCTTACCA 727
Qy 531 --ACGACGCGAGCGGCAACAAACACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTC 588
Db 728 GTATGACAGTGGCGAGATTCTCGGGGATACATCCCATCAGTCACTAAACAGCAGATCTC 787
Qy 589 CGTGGCGGTGAGCGGTGAGTTTCC--GATTACCAACCGGATATGCCACCGGAGGTCTCGT 647

Db 788 AGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGCTCCCACTGGGAACATTGG 847
Qy 648 CTTCCCTGCACAAGGCGCATCCATCCGTACCCCTGTGTCTGAGCAGGAGTGCCAAAGTGGG 707
Db 848 CAAGCGTGTCCATTTC-CCTATGTGAACCAATTCGCCAAAATCCGTCAAGGGAGTTCTCTG 906
Qy 708 ACGAGAAGAAAGAACTGAGCTGGAGGAGAGATGAGACAGCTGGAAGTCCAGCAGGGCA 767
Db 907 GTAGCATTTGGGAATGTTCCCTGGAAAGAGAGGGTTGATGGCTGGAAAAATGAACGAGSACA 966
Qy 768 -----TCC 770
Db 967 AGGGACGATTTCCCATGACGAATGSCACAAAGCAATTTGCTCCCTCTGAGGGTGGGTG 1026
Qy 771 TCGGGCGGGCGCCGATCCCGAAGACATGAGCGCGACGTGGCACTGAACGACGAGGGCA 830
Db 1027 GTGATATTGATGATCACTAATGATTACAAATGGAAGATGCTTTATTGAACGAGCAAACTC 1086
Qy 831 GGCAGCCGCTGTCGAGGAAGGTGTCGATCGGTGAGCAAGGTGAACCGTACCGGATGG 890
Db 1087 GACAGCCTTATCTAGSAAAGTTCCACTTCTCTCCAGGATAAATCCATACAGGATGG 1146
Qy 891 TGATCGTGTGGTCTCGTGTGCTCGCTTCTTCTCCGCTACCGTATCCGTGACACCCG 950
Db 1147 TCATTGTGCTGGATTGATTGTTCTAAGCATCTTCTTGCACTACCGTATCAAAATCCTG 1206
Qy 951 TCCCGAGCCCATCGGGCTGTGGCTGCTCTCATCATCTGCGAGATCTGGTTCGCACT 1010
Db 1207 TGGCAATGCAATACCAATATGCTTCTATCTGTTATATGTGAGATCTGGTTGCTCTT 1266
Qy 1011 CTTGATCCTGACACAGTTCCCAAGTGTTCCTCCCATCGACCGGAGAGAGTACCTCGACC 1070
Db 1267 CGTGGATATTGATCAGTTCCCTAAGTGTTCCTAATCAACCGGGAGAGCTACCTTGATA 1326
Qy 1071 GCCTCTCCTCAGGTACGAGAGGGAAGGGAGCGTCTGCTGTGCGCGGTGACCTGT 1130
Db 1327 GCGTGCATTAAAGGTATGACCGGGAAGGTGAGCACTCAGTTGCTGCTGTGACATTT 1386
Qy 1131 TCGTGAGACGGTGACCCGCTCAAGGAGCGCGCTGTGTGACCGCCCAACACCGTGTCT 1190
Db 1387 TCGTCAGTACAGTGACCCCAATGAAGGAGCTCTCTTGTCACTGCCAATACCGTGTAT 1446
Qy 1191 CATCTCGCGTAGACTACCCGTTGGAAGAGTCTCTCTGTCTGCTACGTCCTCGACGAGCGG 1250
Db 1447 CCATTCTTGTGTGATTACCCGTGTGGATAAGTCTCTTGTCTATGATATCTGATGATGAG 1506
Qy 1251 CGTGCATGCTGACGTTCAAGTTCGCTGCGAGAGCGCGAGTTGCGCGCAAGTGGGTGC 1310
Db 1507 CTGCAATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTCTAGAAAAATGGGTAC 1566
Qy 1311 CTTCTGCAAGAAGTTTCGGCATCGAGCCCGCGCCCGGAGTTCTTACTTCTCGCTCAAGG 1370
Db 1567 CATTTGTTAAGAAGTACAACATTGAACCTAGAGCTCTCTGAATGGTACTTCTCCAGAAA 1626
Qy 1371 TCGACTACTCTAAGACAAAGTTGACGCCACCTCTGTGAGGAGCGCGCGCATGAAGA 1430
Db 1627 TTGATTACTTGAAGACAAAGTGCACCTTCAATTTTAAAGACCGCGCGGCCATGAAGA 1686
Qy 1431 GAGAGTATGAGGAGTTCAAGGTTCGGATCAACGCGCTGTGCGCAAGGCCATGAAGTGC 1490
Db 1687 GAGAAATATGAAGANTTAAAGTTAGGTTAAATGCGCTTTGTGCTAAGGCACAGAAATTC 1746
Qy 1491 CGGACAGGGGTGGATCATAGAGGACGGCAGCGCGTGGCCCGGGAAACACACCCCGACC 1550
Db 1747 CTGAGNAGGATGGATCATGCAAGATGGCACACCATGCGCAGGAACAAATACGAGGACC 1806
Qy 1551 ACCCGGATGATCAGGTGTTCTTGGGCGCAGCGCGGCGCACGACACCGAGGGCAACG 1610
Db 1807 ATCTCTGGAATGATTACAGGTTTCTTGTGTCAAGTGTGGCTTGATCTGAGGCAATG 1866
Qy 1611 AGCTCCCGCTCTGTTAGCTCTCCCGTGAAGAGCGCGCGGATTCAGCACCAACAGA 1670
Db 1867 AGCTACCCCGTTTGGTCTATGTTTCTCGTGAAGAGCGCTCTGGAATTCAGCATCAAGA 1926

Qy 1671 AGCCGCGGCCCATGAACGCTCTGATTCGCGTCTCCGCGTGTGACCAACCGCGCATTTCA 1730
Db 1927 AAGCTGGTGCCATGAATGCTCTTGTGTGTGTCTCAGCTGTCTACCAATGCAACAATACA 1986
Qy 1731 TGCTCAACTTGGACTGTGATCAGCTACATCAACACAGCAGGCGCATCCGGAGGCCATGT 1790
Db 1987 TGTGAAATCTTGAATGTGATCAGTACATTAACAACAGTAAGGCTCTCAGGAAAGCTATGT 2046
Qy 1791 GCTTCTCTCATGAGACCTCAGGTCCGCGGAGAGTCTGCTACGTTTCAAGTCTCCCGCAGAGGT 1850
Db 2047 GCTTCTTATGAGACCTTAACTAGAGAGAGTGTCTGCTACGTCCAGTTTCCCGCAGAGAT 2106
Qy 1851 TCGACGGCATCGACGTGCACGACCGATACGCTAAACAGAAACACCGTCTTCTTCGACATCA 1910
Db 2107 TCGATGCAATTTGACAGGAATGATCGATATGCAACAGSAAACACCGTGTGTTTTCGATATTA 2166
Qy 1911 ACATGAAGGGCTCGAGCGGCATCCAAGSCCGGTGTAGCTCGGACAGGGTGCCTGTTC 1970
Db 2167 ACTTGAGAGGTCTTTGATGGCATCCAAGACAGTTTATGTCCGAACTGGCTGTGTTTTCA 2226
Qy 1971 GCGCCAGGGCGCTCTACGGCTTACAAACCTCTCCAAAGGGACCCCAAGA----- 2015
Db 2227 ACCGAAACAGCTCTATATGTTTATGAGCCCCCAATTAAGCAGAAAGAGGTGTTTCTTGT 2286
Qy 2016 -----GGCCCCAAGATGTGACCTGCGACTGCTGCCCTGTGCTCGGCCGCAAGA 2063
Db 2287 CATCACTATGTGGCGGTAGGAAGAGGCAAGCAAAATCAAAGAGGGCTCGGACCAAGAAGA 2346
Qy 2064 AGCGGAACACGCCAAGGAC-----GGGCTGCCGGAGG 2096
Db 2347 AGTCGCAAGAGCATGTGGACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGG 2406
Qy 2097 GCACCGCTGATATCGGAGTAGATAGCGACAGGAGAGTGTCTATGTCCTCCACATGAACCTTCG 2156
Db 2407 GAGTTGAAGCGCTGGATTTGACGACGAGAAATCACTTCTATGTCTCAATAGAGCTGG 2466
Qy 2157 AGAAGCGGTTGCGGACGATCCCGCGGCTTGTGTCACGTCGACCTGATGGAGGAAGCGCGG 2216
Db 2467 AGAAGAGATTGGCCAGTCCCGACGTTTGTGCTCCACTCTCTGATGGAGTATGGTGTG 2526
Qy 2217 TCCCTCTCTGTCAGCCCCCGCTCTCAAGAGGCGCATCCATGTCTATCAGCTCGG 2276
Db 2527 TTTCTCAGTCCGCAACTCCGAGGTCTCTTCTGAAGAAGTATCCATGTTTATAAGCTGTG 2586
Qy 2277 GCTACGAGGACAAACACGACTGGGGCTGGAGCTGGGCTGGATCTACGGGTCTGATCAAGG 2336
Db 2587 GCTATGAGACAAAGCTGAATGGGGAATGAGATCGGGTGGATCTACGGTTCTGTGACAG 2646
Qy 2337 AGGACATCTGACGGGTTCAAGATGCACTGCGCGGGTGGCGTCCGTTGATCTGCAATGC 2396
Db 2647 AAGACATTTCTCACCGGATTCAGATGCAACGCGGAGGCTGGCGGTCCGATCTACTGCAATGC 2706
Qy 2397 CGAAGCGGGCGGCTTCAAGGGGTGGGCGCGATCAATCTATCGACGCTCTCAACAGG 2456
Db 2707 CCAAGCGGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTTCGACCGCTCTGAACCCAGG 2766
Qy 2457 TGCTCCGTTGGGCGCTCGGGTCTCGAGATCTTCTTTCAGCGGCAAGACCCCTGCTGT 2516
Db 2767 TGCTCCGTTGGGCTTTGGGTCCTGGAGATCTCTTTCAGCGCGGCACTGCCCTGTGTGT 2826
Qy 2517 ACGGCTTCAAGAAACGGCAACCTCAAGTGGCTGGAGCGGTTGCGCTTACATCAACACCA 2576
Db 2827 ACGGCTAC--GGAGGGCGGCTCAAGTTCTTGGAGAGATTGCGGTATCAATCAACACCA 2883
Qy 2577 TCTACCCCTTCACTCTCGCTCCGCTGCTCGCTACTGCACTCCCGCTCCCGCTGCTGCCCTCC 2636
Db 2884 TCTACCCGCTCACTGCTCCATCCGCTTCTCATCTACTGATCTTCCCGGCACTGTGTGTC 2943
Qy 2637 TCACCGCAAGTTTCACTGCGGTGATTAAGAGGTTTGGAGGCTTCCGCTTACATCAACACCA 2696
Db 2944 TCACCGGAAGTTTCACTTCCAGAGATCAGCAACTTCGCGAGGATCTGGTTCATCTCCC 3003

QY 1131 TCGTCAGCAGCGTGGACCCGCTCAAGGAGCCCGCTGGTGACCGCCACACCGGTCTCT 1190
Db 1387 TCGTCAGTACAGTCGACCCCAATGAAGAGCCCTCTCTTGTCTACATGCCAATACCGGTCTAT 1446
QY 1191 CCATCCTCGCGGTAGACTACCCCGGTGGACAAGGTCTCTGCTTACCTGCTCCGACGACGCG 1250
Db 1447 CCAITCTTGTCTGTGATTAACCTCTGTGGATAGGTCTCTTGTCTATGTAATCTGATGAG 1506
QY 1251 CGTCAGTCGTGACGTTGAGTCTGTGCGGAGACGGCCGAGTTCGCGCGCAAGTGGGTGC 1310
Db 1507 CTGCGATGCTGACATTTGATGCACTAGCTGAGACTTCAAGAGTTTCTAGAGTTTCTAGAAAATGGGTAC 1566
QY 1311 CCTTCTGCAAGAAGTTTCGGCATCGAGCCCGCCCGCGGAGTTTCTACTTCTCGTCAAG 1370
Db 1567 CATTTGTTAAGAAGTACAACTTGAACCTTAGAGCTCTGAAATGGTACTTCTCCCAAGAAA 1626
QY 1371 TCGATACCTCAAGACAAAGTTCGAGCCCACTTCGTGACGAGCGCGCGCCATGAAGA 1430
Db 1627 TTGATTACTTGAAGGACAAAGTGACCCCTCATTTGTTAAGACCGCCGGGCCATGAAGA 1686
QY 1431 GAGAGTATGAGAGTTCAAGGTTCGGATCAACGCGCTGTGGCCCAAGGCCATGAAGTGC 1490
Db 1687 GAGAATATGAAGAAATCAAGGTTAGGGTAAATGGCTTTGTTGCTAAGGCACAGAAAGTTC 1746
QY 1491 CGGCAGAGGGTGGATCATGAAGACGGACCGCCGTGGCCCGGGNAACAACCCCGAC 1550
Db 1747 CTGAGGAAGGATGGATCATGAAGATGGACACCATGGCCAGGAAACAATACGAGGAC 1806
QY 1551 ACCCGGATGATCAGAGTGTTCCTGGCCACAGCGCGGCCACGACACCGAGGCAACG 1610
Db 1807 ATCTGGAAATGATTCAGGTTTCTTGTGTCAAGTGTGGCTTGATGTAAGGCAATG 1866
QY 1611 AGTGTCCCGCTCGTGTACGTTCCCGTGAAAGCGCCCGGATTCAGACCAACAAGA 1670
Db 1867 AGCTACCCCGTTTGGTCTATGTTCTCGTGAAGAGCGTCTTGATTCAGCATCAAGA 1926
QY 1671 AGCCCGGCGCATGACGCTCTGATTCGGCTCTCGCCGTGCTGACCAAGCGCCATTC 1730
Db 1927 AAGCTGGTGCATGAATGCTCTTGTGTCTCAGCTGTGCTTACCAATGGACAATACA 1986
QY 1731 TGCTCAACTTGGACTGTGATCACTACATCAACAAGCAAGCCCATCGGAGGCGCATGT 1790
Db 1987 TGTGAACTTGTGATCACTACATTAACACAGTAAAGCTCTAGAGGAGCTATGT 2046
QY 1791 GCTTCTCATGACCCCTCAGTTCGCGCGAAGGTCTGCTAGTTTCACTTCCCGCAGAGGT 1850
Db 2047 GCTTCTTATGGAACCTAACTAGGAAGAGTGTCTGCTAGTCCAGTTTCCCGCAGAGAT 2106
QY 1851 TCGACGGCATCGAGTCGACGACCGATAGCTTAACAGGAACAACCGTCTTCTTCGACATCA 1910
Db 2107 TCGATGGCAATGACAGGAATGATCGATATGCCAACAGGAACACCGGTGTTTTTCGATATTA 2166
QY 1911 ACATGAAGGGCTGGAAGCGCATCAAGGCCGCTGTAGCTCGGACAGGGTGGTGTTC 1970
Db 2167 ACTTGAGAGTCTTGATGGCATCCAGAACAGTTTATGTGGAACCTGGCTGTGTTTCA 2226
QY 1971 GCGCCGAGGCGCTCTACGGCTACAACTCCCAAGGACCCCAAGA----- 2015
Db 2227 ACCGAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAGAGAGGGTGTGTTCTTGT 2286
QY 2016 -----GGCCCCAAGATGTGACCTCGGACTGTGCTCGCCGTGCTTCGGCCGCAAGA 2063
Db 2287 CATCACTATGTGGCGGTAGGAAGAAGGCAAGCAATCAAGAGGGCTCGGACAAAGA 2346
QY 2064 AGCGGAACACGCCAAGAC-----GGGCTGCGCGAGG 2096
Db 2347 AGTCGAGAGCATGTGACAGATTCTGTGCCAGTATTCACTTGAAGATATAGAGGAGG 2406
QY 2097 GCAACGCTGATATGGAGTAGATAGCCAGAGGAGATGCTCATGTGCCACATGAATTCG 2156
Db 2407 GAGTTGAAGGCGCTGGAATTTGACGACGAGAAATCACTTCTTATGTCTCAATATGACCTGG 2466
QY 2157 AGAAGCGGTTCCGGCAGTCCCGCGGCTTCGTCACTGTCGACGCTGATGGAGGAAGCGGCG 2216

Db 2467 AGAAGAGATTTGGCCAGTCCCGCAGCGTTTGTTCCTCCACTCTGATGAGTATGGTGGTG 2526
QY 2217 TCCCTCTCTTCGTTCGAGCCCCCGCGCTCTCAAGGAGGCCATCCATCTCATCACTCGG 2276
Db 2527 TTTCTCAGTCCGCAACTCCGAGTCTCTTCTGAAAGAAGCTATCCATGTTATAGCTGTG 2586
QY 2277 GCTACGAGGACAAGACCGACCTGGGGGTGGAGCTGGGGTGGATCTACGGGTTCGATCA 2336
Db 2587 GCTATGAGGACAAGACTGAATGGGGAACCTGAGATCGGGTGGATCTTACGCTTCTGTGACAG 2646
QY 2337 AGGACATCTCTGACGGGGTTCAAGATGCATCTCCGCGGGTGGCTCCGCTGACTGCATGC 2396
Db 2647 AAGACATTTCTCACCGGATTTCAAGATGCACGCGGAGGTGGCGGTGCGATCTACTGCATGC 2706
QY 2397 CGAAGCGGCGCGCTTCAAGGGGTTCGCGCGCGATCAATCTATCGGACCGCTCAACACGAG 2456
Db 2707 CCAAGCGGCGGACTTTCAAGGGGTTCGCGCGCGATCAATCTTTTCGACCGCTCTGAACGAG 2766
QY 2457 TGCTCCGGTGGGCGCTGGGGTCCGTGAGATCTTCTTCAGCGGGGACAGCCCGCTGTCTGT 2516
Db 2767 TGCTCCGGTGGGCTTTCGGGTCCGTGGAGATCTCTTCAGCGGGGCACTGCGCCCTGTGT 2826
QY 2517 ACGGCTACAGAAACGCGCAACCTCAAGTGGCTGGAGCGCTTCGCTACATCAACACCA 2576
Db 2827 ACGGCTAC---GAGAGGGCGGCTCAAGTTCTTGGAGAGATTCGCGTACATCAACACCA 2883
QY 2577 TCTACCCCTTCACTCCGCTCGCTCCGCTGCTGCGCTACTGCACCTCCCGCGCTCTGCTCTCC 2636
Db 2884 TCTACCGCTCAGCTCCATCCGCTTCTCATCTACTGATCTCTGCGCGCATCTGCTCTGC 2943
QY 2637 TCACCGGAAGTTCAATCATGCGCTGATAGCAGTTCGCGAGGCTCTTCTTCACTGCGCC 2696
Db 2944 TCACCGGAAGTTCAATCATTCAGAGATCAGCAACTTCGCGAGCATCTGTTTCACTCTCC 3003
QY 2697 TCTTATGTTCCATCTTCGCGAGCGGCATCTGGAGATCGCGTGGAGCGGGTGGGATCG 2756
Db 3004 TCTTATGTTCCATCTTCGCGAGCGGCATCTGGAGATCGCGTGGAGCGGGTGGGATCG 3063
QY 2757 AGGAGTGTGGAGGAAGCAGCAGTTCTGGGTTCATCGCGCGCTGTCTCGCGCATCTCTTTCG 2816
Db 3064 AGGAGTGTGGAGGAACGACGAGTTCTGGGTTCATCGGGGATCTCTCGCGCATCTTTCG 3123
QY 2817 CCGTGTGCGAGGCGCTGTCAAGGTCTTCGCGCGGATCGACAACAACTTCAACGCTCACT 2876
Db 3124 CCGTGTTCAGGGCTGTCTCAAGGTGTGGCGCGCATCGACAACAACTTCAACGCTCACT 3183
QY 2877 CCAAGGCAACGCGGACGAGGACGAGTTTCGCGGAGCTTACGCTTCAAGTGGACCA 2936
Db 3184 CCAAGGCTCG---GACGAGGACGCGCATCTTCGCGAGCTGTATGTTCAAGTGGACGA 3240
QY 2937 CGCTCTCTATCCCGCCACCACTGCTCATTAACGTCATCGGCTCGTGGCGCGCA 2996
Db 3241 CGCTCTGTATCCCGCCACCACTGCTCATTAACGTCATCGGCTCGGCTGTCTCGCGGCA 3300
QY 2997 TCTCCGACCGCATCAACACAGGGTACAGTCTTCGGGGCGCTCTTCGGCAAGCTCTTCT 3056
Db 3301 TCTCTACGCGCATCAACAGCGGATACAGTCTGGGGCGCTCTTCGGCAAGCTCTTCT 3360
QY 3057 TCGCTTCTGGGTATCGTCACTCTAAGGTTCCTCAAGGGGCTCATGGGGCGCCAGA 3116
Db 3361 TCGCTTCTGGGTATCGTCACTCTAAGGTTCCTCAAGGGGCTCATGGGGCGCCAGA 3420
QY 3117 ACAGGACCGCCACCGTGTGTCATCTGCTTCTGCTGCTGCTCATCTTCTCTCCCTGC 3176
Db 3421 ACAGGACCGCCACCGTGTGTCATCTGCTGCTGCTGCTGCTGCTCATCTTCTCTCTTCG 3480
QY 3177 TCTGGGTGAGATCGACCTTTTCATGCTCAGGACCAAGGGCGCCGAGCTCAGGAGTGTG 3236
Db 3481 TGTGGGTGCGATCGACCCCTTACCAACCGCGTCACTGGCGCGGATACCCAGAGCTGTG 3540
QY 3237 GCATCAATGCT 3248
|||||

Db 3541 GCATCAACTGCT 3552

RESULT 13
US-10-209-059-9
: Sequence 9, Application US/10209059
: Publication No. US20030163838A1
: GENERAL INFORMATION:
: APPLICANT: Dhugga, Haiyin
: APPLICANT: Wang, Haiyin
: TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
: FILE REFERENCE: 0864R2
: CURRENT APPLICATION NUMBER: US/10/209,059
: PRIOR FILING DATE: 2002-07-31
: PRIOR APPLICATION NUMBER: 60/096,822
: PRIOR FILING DATE: 1998-08-17
: PRIOR APPLICATION NUMBER: 09/371,383
: PRIOR FILING DATE: 1999-08-06
: PRIOR APPLICATION NUMBER: 09/550,483
: PRIOR FILING DATE: 2000-04-14
: NUMBER OF SEQ ID NOS: 52
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 3773
: TYPE: DNA
: ORGANISM: Zea mays
US-10-209-059-9

Query Match 40.68; Score 1397.6; DB 16; Length 3773;
Best Local Similarity 67.2; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;

QY 133 GAACGAGCTGCTGTGATCCGGGCGCACGAGGACCCCAAGCGCTGGGGCGCTGAGCGG 192
DB 325 GAGTCGCTTGCCATGAGGCGGACGCGGACGGGTGAGTTCGGGAGGCGCGGTGGCGG 384
QY 193 GCAGGTGTGCGAGATATGCGCGGACGAGGTGCGGCTCACGGTGGAGCGGACCTCTTCGT 252
DB 385 ACAGGTGTGCGAGATCTGCGGCGACGCGGTGGGACCCACGCGGAGGGGACGCTTCGCG 444
QY 253 CGCTGTCAACGAGTGGCGCTTCCCGGTGTCGGGCGCTGTACAGGTACGAGCGCGGGA 312
DB 445 CGCTGTGCGAGCTGTGCGGGTTTCGGGTGTCCGCGCTGTACGAGTACGAGCGCAAGGA 504
QY 313 GGGCACGCGAACTGCCCGCCAGTGCAGACGCGCTACAAGCGCTCAAGGGGAGCGCGAG 372
DB 505 CGGACGCGCGGTGCGCCCGAGTGCAGACCAAGTACAGCGGCCACAGGGGAGCGCGGC 584
QY 373 GGTTCGCGGACGATGACGAGGAGGACATGCGACGCTTGGAGCAGAGTTCAACATCGA 432
DB 565 GATCCGTGGGGAGGAGGAGACGACACTGATGCCGATAGCGACTTCAATTACCTTGCATC 624
QY 433 CGACGGAATCAGCAGAGGAGCTGGAGGCGCAATCGACATCGACAGCGCAGATCA----- 485
DB 625 TGGCAATGAGGACCGAGAGGAGATTTGCCGACAGAAATTCGCGACAGAAATCGCAGCTGGCGCATGAACGT 684
QY 486 -----CCGAGGCGATCTGCAGCGCAGGATGAGCTACGGGAGGGGCGCG----- 530
DB 685 TGGGGGAGCGGGGATGTTGGTGCCTCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAA 744
QY 531 --ACGACGGGACGGGCAACACACCCCGCAGATCCCGCCCATCATCAACCGGCTCCCGCTC 588
DB 745 GTATGACAGTGGCGAGATTCCTCGGGGATACATCCCATCAGTCACTAACAGCCAGATCTC 804
QY 589 CGTCCCGGTGAGCGGTGAGTTTC--GATTACCAACGGGTATGGCCAGCGGAGGTCTCGT 647
DB 805 AGGAGAAATCCCTGGCTTCCCGCTGACCATCATATGATGTCCCACTCGGGAACATTTGG 864
QY 648 CTTCCCTGCAACAGCGCATCATCCGTAACCTGTGTCTGAGCCAGGAGGTGCCAAGTGGG 707
DB 865 CAAGCGTGTCTCATTTTC--CCTATGTGAACCATTCGCCAAATTCGTCAGGGAGGTCTCTG 923

QY 708 ACGAGAAGAAAGAGTGAAGAGGAGGATGACGACTGGAAGTCAAGCAGGCGCA 767
DB 924 GTAGCAATTGGGAATGTTGCTCGAAAGAGAGGGGTTGATGGCTGGAATAAAGAGCAGGACA 983
QY 768 -----TCC 770
DB 984 AGGGGACGATTCATGACGAATGGCAAGCAATTCCTCTGTAGGGTTCGGGGTGTG 1043
QY 771 TCGGCGGCGCGCGCATCCGAAGACATGGACCCGACGTGGCACTGAACGACGAGGCGCA 830
DB 1044 GTGATATTGATGATCAATCACTGATTACAAATGAAAGATGCTTATTGAACGACGAATCTC 1103
QY 831 GGACGCGCTGTGCGAGAAAGGTGTCGATCGCGTCGAGCAAGGTGAACCCCTACCGGATGG 890
DB 1104 GACAGCTCTCTATCTAGGAAAGTTCACCTTCCTCCAGGATAAATCCATACAGGATGG 1163
QY 891 TGATCTGTGTGCTCTCGTTGTGCTCGCTTCTTCTCCGGTACCGGTATCCTGCACCCCG 950
DB 1164 TCATTGTGTGCGATTTGATTTGTTCAAGCATCTTTCGCACTACCGGTATCACAATCTCTG 1223
QY 951 TCCCGACGCGCATTCGGGCTGTGGCTGTCTCCATCATCTGCGAGATCTGGTTCGCCATCT 1010
DB 1224 TGGCAATGCATACCATTAATGCTTCTATCTGTTATATGTGAGATCTGGTTCCTCTT 1283
QY 1011 CTTGGATCTCGACAGTTCCCAAGTGTTCCTCCATCGACCGCGAGACGTACTCGAC 1070
DB 1284 CGTGGATAATTGGATCAGTTCCCTAAGTGTTCCTCAATCAACCGGAGACGTACTCGATA 1343
QY 1071 GCCTCTCCCTCAGGTACGAGAGGAGGAGCGGCTGCTGTGCTGTGGCGGTGCACTGT 1130
DB 1344 GGTGTCATTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGTGTTGACATTT 1403
QY 1131 TCGTAGACACGCTGAGACCGCTCAAGGAGCGCGCTGTGTGACCGGCAACACCGTGTCT 1190
DB 1404 TCGTAGTACAGTCGACGCCAATGAAGAGGCTCTCTTGTCACTGCCAATACCGTGTAT 1463
QY 1191 CCATCTCGCGTAGACTACCCCGTGGACAAAGTGTCTCTGTACTAGTCTCGACGAGCGG 1250
DB 1464 CCATCTGTGTGGATTAACCTGTGGATAAGGTCTCTTGTATGTATCTGTATGATGAG 1523
QY 1251 CGTGTGATGTGAGTTCGAGTGTGCTGTGCGAGAGCGCGGAGTTTCGCGCAAGTGGTGC 1310
DB 1524 CTGCGATGTGACATTTGATGACACTAGCTGAGACTTTCAGAGTTTGTAGAAAATGGGTAC 1583
QY 1311 CCTTCTGCAAGAGTTTCGCGCATCGAGCGCGCGCGCGGAGTTCTACTTCTCGCTCAAGG 1370
DB 1584 CATTTGTAAGAGTACAACTTGAACCTTAGAGCTCTGATGCTTCTCCAGGAAA 1643
QY 1371 TCGACTACCTCAAGGACAGGTGTGACGCCACCTTCGTGCAAGGAGCGCGCGGCGCATGA 1430
DB 1644 TTGATTTACTTGAAGGACAAAGTGCACCTTTCATTTGTTAAAGACCGCGCGGCGCATGA 1703
QY 1431 GAGAGTATGAGGAGTTCAAGGTCCGGATCAACGCGCTGTGGCGCAAGGCCATCAAGTGC 1490
DB 1704 GAGAAATGAAGAAATTCAAAGTTAGGGTAAATGGCTTGTGCTAAGGCGACAGAAAGTTC 1763
QY 1491 CGGACAGAGGCTGGATCATGAAGGACGCGCGCTGGCGCGGGAACAAACACCGCGGACC 1550
DB 1764 CTGAGGAGGATGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACAGGAGACC 1823
QY 1551 ACCCGCGCATGATTCAGGTGTTCTCGGGCCACAGCGCGGCGCCACGACACGAGGGGCAAG 1610
DB 1824 ATCTGGAATGATTCAGGTGTTCTTGTGTCACAGTGTGGCTTGTACTAGAGGGCAATG 1883
QY 1611 AGCTGCGCGCTCGGTGATGCTCCCGTGAAGGCGCGCGGATTCACACACCAAGA 1670
DB 1884 AGCTACCCGCTTGGTCTATGTTCTCGTGAAGAGCGTCTCGATTCACGATCACAGA 1943
QY 1671 AGCGCGCGCATGAACGCTCTGATTCGGGTCTCCCGCGGTCTGTCGACCAACGCGGCCATTC 1730
DB 1944 AAGCTGTCGTCATGATGCTCTTGTGCTGTCTGCTGCTGTGCTTACCAATGAGCAATACA 2003
QY 1731 TGCTCAACTTGGACTGTGATCATACATCAACAAAGCAAGGCCATCCCGGGAGGCCATGT 1790

Db 2004 TGTGATCTTGAATTTGATACATTAACAACAGTAAGGCTCTCAGGAAGCTATGT 2063
Qy GCTTCTCATGACCCCTCAGGTCCGCGGAAGGTCTGCTAGGTTCCAGTTCCCGAGAGGT 1850
Db GCTTCTTATGGAACCTTAACCTAGGAAGGTGCTGCTAGTCCAGTTCCCGAGAGAT 2123
Qy TCGAGCGCATCGAGTGCAGACCGATACGCTAACAGGAACACCGTCTTCTTCGACATCA 1910
Db TCGATGGCAATTGACAGGAATGATCGATATGCCAACAGGAACACCGTGTTCGATATTA 2183
Qy ACATGAAGGGCTGAGAGCGCATCAAGGCCGCTGCTAGCTCGGACAGGGTGGTGTTC 1970
Db ACTTGAGAGGTCTGATGGCATCCAGGACCAAGTTTATGTCGAACTGGCTGTGTTTCA 2243
Qy GGCGCCAGCGCTCTACGGCTACAAACCTCCCAAGGGAACCAAGA ----- 2015
Db ACCGAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAGAGAAGGGTGTGTTCTTGT 2303
Qy ----- GGGCCAGATGTTGACCTGGACTGGGACTGCTGCCGCTGCTCGGCCGCAAGA 2063
Db CATCATATGTTGGCGGTAGGAAGGCAAGCAATCAAGAGGGCTCGGACAAAGA 2363
Qy AGCGGAACACGCCAAGGAC-----GGGCTGCCGGAGG 2096
Db AGTCGAGAGATGTTGACAGTTCGTGCGCAGTATTCAACCTTGAAGATATAGAGGAG 2423
Qy GCACCGCTGATATGGGAGTAGATAGCGACAAGAGGATGCTCATGTCCACATGAATTCG 2156
Db GAGTTGAAGGCGCTGGAATTCACGACGAGAAATCACTTCTATGTCTCAATAGGCTGG 2483
Qy AGAAGCGTTCCGGGAGTCCCGGGGTTTCGTCACGTCGACGCTGATGAGGAAGAGCGCG 2216
Db AGAAGAGATTTGGCGAGTCCCGAGCGTTTGTGCTCCATCTGATGGAGTATGGTGGTG 2543
Qy TCCCTCTCTGTCGAGCCCCCGCGCTCTCAAGGAGGCCATCATGTCATCAGCTGCG 2276
Db TTCCTCAGTCGCAACTCCGAGTCTCTTCTGAAGAAGCTATCCATGTTTAAAGCTGTG 2603
Qy GCTACGAGGAACAAGCTGAGGCTGGGCTGGAGCTGGGTTGATCTACGGGTCGATCAAG 2336
Db GCTATGAGGAACAAGCTGAATGGGGAACCTGAGATCGGGTGATCTACGGTTCTGTGAC 2663
Qy AGGACATCTGAGCGGGTTCAAGATGCACTCCGCGGGTGGCGTCCGTTGATGCTATGCT 2396
Db AAGACATTTCTACCGGATTTCAAGATGCAAGCGCGGAGGCTGGCGGTCTGATCTACTG 2723
Qy CGAAGCGGCGCGTTCAAGGGTCCGCGCGCATCATCTATCGGACCGTCTCAACGAG 2456
Db CCAAGCGGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTTCGGACCGTCTGAACG 2783
Qy TGCTCCGTTGGCGCTGGGGTCCGTCGAGATCTTCTCAGCGCGGCACAGCCCGCTGCTGT 2516
Db TGCTCCGTTGGCTCTTGGTCCGTCGAGATCTTCTCAGCGCGCATGCTGCCCTGTGCT 2843
Qy ACGGCTACAGAAACGCAACCTCAAGTGGCTGAGCGCTTCGCTTACATCAACCAACCA 2576
Db ACGGCTAC--GAGAGGGCGGCTCAAGTTTCTGGAGAGATTGCGGTACATCAACCAACCA 2900
Qy TCTACCCCTTCACTCGCTCCGCTGCTGGCTACTGACCCCTCCCGCGCTCTGCTCC 2636
Db TCTACCCGCTCAGCTTCATCCGCTTCTCATCTACTGATCCTGCCCGCCATCTGTCTG 2960
Qy TCACCGGAAGTTTATGATTCGATTCAGCATGTTTCGCGAGCTCTTCTTCACTCGCC 2696
Db TCACCGGAAGTTTATGATTCGAGATTCAGCAACTTCGCGAGCATCTGTTTCACTCC 3020
Qy TCTTCACTGCTCATCTTCGCAAGCGGATCTTGGAGATCGCGTGGAGCGGGTGGAGCAT 2756
Db TCTTCACTGCTCATCTTCGCAAGCGGATCTTGGAGATCGCGTGGAGCGGGTGGAGCAT 3080
Qy AGGAGTGGTGGAGGAACAGAGAGTTCGGGTTCATCGCGCGGCTGTCGCGCATCTCTTCG 2816

Db 3081 ACGAGTGGTGGAGGAACGAGCAGATTCTGGGGTGATCGGGGGGATCTCCGCGCACCTCTTCG 3140
Qy CCCTCGTGCAGGGGCTGCTCAAGGTCTCTCGCGGGATCGACACCAACTTTCACCGGTCACT 2876
Db CCGTGTTCAGGGGCTGCTCAAGGTCTGCTGGCGGCATCGACACCAACTTTCACCGGTCACT 3200
Qy CCAAGGCCACCGGGCGAGAGGACGAGTTCCGCGGAGCTTACATGTTTCAAGTGGACGA 2936
Db CCAAGGGCTCG--GACGAGGACGGCGACTTTCGCGGAGCTGTACATGTTTCAAGTGGACGA 3257
Qy CGCTCTCATCCCCCGCCACACCGTCTCATCAATTAAGTCAATCGGGTCTGCGCGGCA 2996
Db CGCTCTGATCCCCCGCCACACCAATCTGATCATCAACTGGTGGGCTGCTGCGCGGCA 3317
Qy TCTCCGACGCGCATCAACCAAGGGTACAGTCTCTGGGGGCCCTCTTTCGGCAAGCTCTTCT 3056
Db TCTCTACGGCATCAACAGCGGATACAGTCTGCGGGCCGCTCTTTCGGCAAGCTCTTCT 3377
Qy TCGCTTCTGGGTCTATGCTCCACCTCTACCGTTCTCAAGGGGCTCATGCGGGCGCAGA 3116
Db TCGCTTCTGGGTCTATGCTCCACCTCTACCGTTCTCAAGGGGCTCATGCGGGCGCAGA 3437
Qy ACAGACGCGCCACCGTTGTTGTCATCTGCTGCTCATCTGCTGGCTCCATCTTCTCCCTGC 3176
Db ACCGACCGCGACCATGCTGCTGCTGCGGCACTCTGCTGGCTCCATCTTCTTCTTCTG 3497
Qy TCTGGGTGAGATACGACCTTTTCATGCTCAGGACCAAGGGGCCGAGTCAAGGAGTGTG 3236
Db TGTGGTTCGATCGACCCCTTTCACCGCGCTCACTGGCGCGGATACCCAGACGTTGTG 3557
Qy GCATCAATTGCT 3248
Db GCATCAACTGCT 3569

RESULT 14
US-10-160-719-5
; Sequence 5, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338) ... (3568)
US-10-160-719-5

Query Match 40.6%; Score 1397.6; DB 16; Length 3773;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;
Qy 133 GAACGAGCTGCTGCTGATTCGCGGGCCACGAGGACCCCAAGCCGCTGCGGGCGCTGAGCGG 192
Db 325 GAGCTCGCTTGGCATGGAGGGCGGACGGGTGAAGTTCGGGGAGGCGCGGTGGCGG 384
Qy 193 GCAGGTGTGAGAGATATGCGCGGACGAGGTCGGGCTCACGGTGGACCGGACCTCTTCGT 252

||||| 385 ACAGGTGTCAGATCTGCGCGACGCGTGGGCACCAACGCGCGAGGGGACGCTTTCGC 444
QY 253 CGCCTGCAACGAGTGGCGCTTCCCGCTGTGCGCGCCCTGTCTACGAGTACGAGCGCGCGGA 312
Db 445 CGCCTGCGAGCTGTCGGGGTTTCGGTGTGCGCGCCCTGTCTACGAGTACGAGCGCAAGGA 504
QY 313 GGGCAGCAGAACTGCCCCCAGTGCACAGACGCGCTACAAGCGCTCAAGGGGAGCCCGAG 372
Db 505 CGGCAGCAGCGCGTGCCTCCAGTGCACAGCAACAGTACAAGCGCCACAAGGGGAGCCCGGC 564
QY 373 GGTTCGCGGGACGATGACGAGGAGGACATGACGACCTGGAGCAAGATTCAACATCGA 432
Db 565 GATCCGTGGGGAGGAAGAGACGACACTGATGCCGATAGCGACTTCAATTACCTTGCATC 624
QY 433 CGACGAGAACTCAGCAGAGGCGAGCTGGAGGGCAACATGCAGAACAGCCAGACATCA----- 485
Db 625 TGGCAATGAGGACCAAGAAAGCAAGATTGGCGACAGAAATCGCGACGTGGCGCATGAACGT 684
QY 486 -----CCGAGGCGATGCTGCAACGCGAGGATGAGCTACGGGAGGGGCCCG----- 530
Db 685 TGGGGGCGCGGGGATGTTGTCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAA 744
QY 531 --ACGACGGCGACGGCAACACACCCCGCAGATCCCGCCCATCATACCGGCTCCCGCTC 588
Db 745 GTATGACAGTGGCGAGATTCCCTCGGGGATACATCCCATCAGTCACTAACACGCGCAGATCTC 804
QY 589 CGTGCCGCTGAGCGGTGAGTTTCC-GATTACCAACGGGTATGGCCACGGCGAGGTCTCGT 647
Db 805 AGGAAATCCCTGGTCTCTCCCTGACCATCATATGATGTCCCAACTCGGAACATTTG 864
QY 648 CTTTCCCTGCAACAGCGCATCCATCCGTACCTGTGTCTGACCCAGGGAGTGCCAAAGTGGG 707
Db 865 CAAGCGTCTCCATTTC-CCTATGTGAACCATTCGCCAATCCGTCAAGGGAGTCTCTG 923
QY 708 ACGAAGAAAGAGTGAAGTGGAGGAGGATGGAAGCATGGAAGTCCAAAGGCGCA 767
Db 924 GTAGCATTTGGGAATGTTTGCCTGGAAAGAGGGGTTGATGGCTGGAAATGAAGCAGGACA 983
QY 768 -----TCC 770
Db 984 AGGGACGATTCCCATGACGAATGGCAAGCATTGTCTCCTCTGAGGGTGGGGGTG 1043
QY 771 TCGCGCGCGCGCGATCCGCAAGACATGGAACGCGCAGTGGCACTGAACGACGAGCGCA 830
Db 1044 GTGATATTGATGCATCAACTGATTACAAATGGAAGATGCTTATTGAAACGACGAATC 1103
QY 831 GGCAGCCGCTGTGAGGAAAGTGTGATCGGTGAGCAAGGTGAACCCGTACCGGATGG 890
Db 1104 GACAGCCTCTATCTAGGAAAGTTCCACTTCTCTCCAGGATAAATCCATACAGGATGG 1163
QY 891 TGATCGGTGTGCTCGTGTGCTCGCCTTCTTCTCCGTACCGTATCCTGCACCCCG 950
Db 1164 TCATTGTGTCGATTAATGTTCTAAGCATCTTCTTGCACTACCGTATCAAAATCCTG 1223
QY 951 TCCCGGACGCCATCGGGCTGTGGCTCGTCTCCATCATCTGCGAGATCTGGTTCGCCATCT 1010
Db 1224 TGGCAATGCATACCCATTATGGCTCTCTATCTGTTATATGTGAGATCTGGTTCCTTT 1283
QY 1011 CTTGGATCTCGACCACTTCCCAAGTGGTTCGCCATCGACCGCGGAGACGTACTCGACC 1070
Db 1284 CGTGGATATTGGATTCAGTTCCTCAAGTGGTTTCCAAATCAACCGGGAGACGTACTTGATA 1343
QY 1071 GCCTCTCCCTCAGTACGAGGGAAGGGAGCGCTGCTGTGTCGGCGGTGGACCTGT 1130
Db 1344 GGTGGCAATTAAAGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGTGTTGACATTT 1403
QY 1131 TCGTGAACGCTGGACCCGCTCAAGGAGCGCGCTGGTGAACCGCAACACCGTGTCT 1190
Db 1404 TCGTCACTAGTCGACCCCAATGAAGGAGCTCTCTTGTCACTGCGCAATACCGTGTCTAT 1463
QY 1191 CCATCTCGCGTACATACCCCGTGAACAGGTCTCTGCTAGTCTCTCGACGACGGCG 1250
|||||

Db 1464 CCATTCTTGCTGTGGATTACCCCTGTGGATAAAGGTCTCTTGCTATGTATCTGTGATGGAG 1523
QY 1251 CGTGCTGTGTGACGTTTCGAGTCTGTCGAGACGCGCCGAGTTCGGCGGCAAGTGGGTGC 1310
Db 1524 CTCGGATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTGTAGAAATGGGTAC 1583
QY 1311 CTTTCTGCAAGAAATTTGGCATCGAGCCCGCGCCCGGAGTTCTACTTCTCTGCTCAAGG 1370
Db 1584 CATTTGTTAAGAGTACAACATTGAACCTAGAGCTCTCGAATGGTACTTCTCTCCAGAAA 1643
QY 1371 TCGACTACCTCAAGGCAAGGTGTCAGCCACCTTCTGTCAGAGAGCGCGGCCCATGAAGA 1430
Db 1644 TTGATTACTTTGAAGGCAAAAGTGCACCTTCAATTTGTTAAAGACCCCGCGGCATGAAGA 1703
QY 1431 GAGAGTATGAGGAGTTCAAGGTCGGATCAACCGCTGGTGGCGCAAGGCCATGAAGTGC 1490
Db 1704 GAGNATATGAAGAAATTCAAAGTTAGGTAAATGGCCCTTGTGTAAAGGCACAGAAAGTTC 1763
QY 1491 CGCAGAGGGGTGGATCATGAAGGACGCGACGCGCTGGCCCGGGAACAACACCCGCGAAC 1550
Db 1764 CTGAGGAAGGATGGATCATGTCAAGATGGCACACCATGGCCAGGAAACAATACCAGGAC 1823
QY 1551 ACCCGCGCATGATCCAGGTGTTCTCGGCCACAGCGCGGCCACGACACGAGGGGCAAG 1610
Db 1824 ATCCTGGAATGATTCAAGGTTTTCCTTGGTCAACAGTGGTGGCCTTGATCTAGAGGGCAATG 1883
QY 1611 AGCTGCCCGCTCGGTGTACGTCTCCCGTGAAGCGCCCGGATTCAGACACCAACAAGA 1670
Db 1884 AGCTACCCGTTTGGTCTATGTTTCTCGTGAAGAGCTCTCGATTTCCAGTATCAACAAGA 1943
QY 1671 AGCCCGCGCCATGAACGCTCTGATTCGCGTCTCCCGCTGTGTCGACCAACCGCCCAATTC 1730
Db 1944 AAGCTGTGTCATGAATGCTCTTGTTCGTCTCAGCTGTGCTTACCAATGGACAATACA 2003
QY 1731 TGCTCAACTTGGACTGTGATCACTACATCAACAGCAGAGGCCATCCGCGGAGCCATGT 1790
Db 2004 TGTGGAATCTTGATTTGTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGT 2063
QY 1791 GCTTCTCATGGACCTCAGGTCGCGGGAAGGTCTGTACGTTCAAGTTCGCCGACAGGT 1850
Db 2064 GCTTCTTATGACCTTACCTAGGAAGGATGTCTGTACGTCAGTTTCCCGCAGAT 2123
QY 1851 TCGACCGCATTCGACGTGCGACGATACGTTAAACAGGAACACCGTCTTCTTCGACATCA 1910
Db 2124 TCGATGGCATTCAGAGGAATGATCGATATGCCAACAGGAACACCGTGTCTTCGATATTA 2183
QY 1911 ACATGAAGGGCTGGAGCGCATCCAGGCCGCTGTACGTGCGGACAGGGTGGTGTTC 1970
Db 2184 ACTTGAGAGTCTTGATGGCATCCAGGACCAAGTTTATGTGCGAACTGGCTGTGTTTCA 2243
QY 1971 GCGCGCAGGCGCTCTACGGCTACAACCTCCCAAGGACCCAAGA----- 2015
Db 2244 ACCGAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAGAGAGGGTGGTTTCTTGT 2303
QY 2016 -----GGCCCAAGATGGTGAACCTGCGACTGTGTCGCCGTCTTCGCGCGCAAGA 2063
Db 2304 CATCACTATGTGGCGGTAGGAAGCAAGCAAAATCAAAAGAGGGCTCGGACAAAGAAGA 2363
QY 2064 AGGGGAAACAGCCCAAGGAC-----GGCTGCCGAGG 2096
Db 2364 AGTCGAGAGCATGTGGACAGTTCTGTGCCAGTATTCACCTTTGAAGATATAGAGGAGG 2423
QY 2097 GCACCGCTGATATGGGAGTAGATAGCGACAAGGAGATGCTCATGTCCCATGAACTTCG 2156
Db 2424 GAGTTGAAGGGCTGATTTGACGACGAGAAATCACTTCTTATGTCTCAATGAGCCTGG 2483
QY 2157 AGAAGCGGTTCCGGCAGTCCGCGGCTGTGTCACGTGCGAGCTGTGAGGAGGAAGGGCG 2216
Db 2484 AGAAGAGATTGGCCAGTCCGACGCTTGTTCCTCCACTCTGATGGAGTATGGTGGTG 2543
QY 2217 TCCCTCTCTGTCGAGCCCGCGCGCTCTCAAGGAGGCCATCCATGTTCATCAGCTGCG 2276
Db 2544 TTCTCAGTCCGCAACTCCGGAGTCTCTCTGAAAGAGCTATCCATGTTTAAAGCTGTG 2603
|||||

```
QY 2277 GCTACGAGACAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGTTCGATCAGG 2336
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2604 GCTATGAGACAAGACTGAATGGGAACCTGAGATCGGGTGGATCTACGGTCTGTGACAG 2663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2337 AGGACATCTGACGGGTTCAAGATGACACTGCGGGGTGGCGCTCCGTGTACTCATGCG 2396
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2664 AAGACATCTACCGGTTCAAGATGACAGCGCGGAGCTGGCGGTGATCTACTGATGCG 2723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2397 CGAAGCGGCGCGTTCAAGGGGTGGCGCGCATCAATCTATCGGACCGTCTCAACCCAGG 2456
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2724 CCAAGCGGCGAGCTTCAAGGGGTCTGCCGCCATCAATCTTTTCGACCGCTTGAACCCAGG 2783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2457 TGCTCGGTGGGCGCTGGGTCGCTCGAGATCTTTTTCAGCCGGACAGCCCGCTGCTGT 2516
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2784 TGCTCGGTGGGCTCTGGGTCGCTGGAGATCTTTTTCAGCCGGACAGCTGCCCGCTGTGT 2843
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2517 ACGGCTTACAAGAAACGCAACCTCAAGTGGCTGGAGCGCTTCCCTTACATCAACACCA 2576
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2844 ACGGCTAC---GGAGGGCGGCTCAAGTTCCTGGAGAGATTCCGTTACATCAACACCA 2900
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2577 TCTACCCCTTACCTCGCTCCGCTGCTGCGCTACTGCAACCTCCCGCGCTCTGCCCTCC 2636
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2901 TCTACCCGCTCAGTCCATCCGCTTCTCATCTACTGATCTGCGGCCATCTGCTGCG 2960
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2637 TCACCGGCAAGTTCAATCATGCCGTGCATGACAGCTTGCAGCGCTCTTTTCATCGGCC 2696
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2961 TCACCGGAAGTTCAATCTCCAGAGATCAGCAACTTGCACGATCTGGTTTCATCTGCC 3020
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2697 TCTTATGTCATCTTCGAGCGGCATCTCGAGATGCGGTGGAGCGGGGTGAGCATCG 2756
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3021 TCTTATCTGATCTTCGCGCAGCGGCATCTCGAGATGAGGTGGAGCGGGGTGAGCATCG 3080
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2757 AGGAGTGTGGAGGAACAGCAGTCTCGGTTCATCGGGCGCTGTCCGCGCATCTCTTCG 2816
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3081 ACGAGTGTGGAGGAACAGCAGTCTCGGTTCATCGGGCGCATCTCCGCGCATCTTCG 3140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2817 CGTCTGTGAGGGCTGTCTAAGGTCTCGCGGGATCGACCAACTTCACCGTCACT 2876
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3141 CGGTCTTTCAGGGCTGTCTAAGGTCTCGCGGGATCGACCAACTTCACCGTCACT 3200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2877 CCAAGGCCACCGGACGAGGACGAGTTCCGCGGCTCTACGCTTCAAGTGGACCA 2936
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3201 CCAAGGCTCG---GACGAGGACGGGACTTCGCGGAGCTGTACATGTTCAAGTGGACGA 3257
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2937 CGCTCTCATCCGCGCCACACGCTGCTCATCATTAAGCTCATCGCGCTCGTGGCGGCA 2996
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3258 CGCTCTGATCCGCGCCACACCATCTGATCATCAACTGCTGGTGGGCTGTGCGCGGCA 3317
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2997 TCTCGGACCGCATCAACAAACGGGTACAGTCTCGGGGCGCCCTCTTCGGCAAGCTCTTCT 3056
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3318 TCTCTACGCCATCAACAGCGGATACAGTCTGTGGGGCGCGCTCTTCGGCAAGCTCTTCT 3377
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3057 TCGCCTTCTGGGTTCATCGTCCACCTCTACCGTTCCTCAAGGGGCTCATGGGGCGCAGA 3116
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3378 TCGCCTTCTGGGTTCATCGTCCACCTGTACCGGTTCTCAAGGGGCTCATGGGGCGCAGA 3437
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3117 ACAGGACGCCACCGTGTGTGTCATCTGCTGCTCATCTGCTGGCGCTTCATCTTCTCCCTGC 3176
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3438 ACCGACCCCGACCATCTGCTGCTGCTGGGCGATCTGCTGGGCTTCATCTTCTTCTTTCG 3497
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3177 TCTGGGTGAGGATCGACCTTTTCATCGTCAGGACCAAGGGCGCCGACGCTCAGGCAAGTGTG 3236
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3498 TGTGGGTTCGATCGACCCCTTCACCAACCGGTCACCTGGCGCGGATACCCAGACGCTGTG 3557
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3237 GCATCAATTGCT 3248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3558 GCATCAATTGCT 3569
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 15

US-10-627-132-9

; Sequence 9, Application US/10627132

```
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-9
```

```
Query Match 40.6%; Score 1397.6; DB 17; Length 3773;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 924; Indels 143; Gaps 9;
```

```
QY 133 GAACGAGCTGTGTGATCCCGGGCCACGAGACCCCAAGCGGCTCGCGGCGCTGAGCGG 192
Db 325 GAGCTCGCTTCCCATGGAGGCGACGCGGACCGGTGAAGTCGGGAGCGCGGTGGCGG 384
QY 193 GCAGTGTGCAGATATATCGCGCGACGAGGTCCGGCTCACGGTGGACGCGACCTCTTCT 252
Db 385 ACAGTGTGCCAGATCTCGCGGACGCGGTGGGACCAACGCGGAGGGGAGCTTTCGCG 444
QY 253 CGCTGTCAACGAGTGTGCGCTTCCCGTGTGCGCGCTTGTACGAGTACGAGCGCGCGGA 312
Db 445 CGCTGTGACGCTCTCGCGGTTTCGGGTGTGCGCGCTTGTACGAGTACGAGCGCAAGGA 504
QY 313 GGGCACGCGAAGTCTCCCGCGAGTGCAGACGCGTACAGCGCTCAAGGGGAGCGCCGAG 372
Db 505 CGGCACGCGGCGTGTGCGCGGCTGCAAGTCAAGCGCCCAACGAGTACGAGGAGCGCGCG 564
QY 373 GGTTCGCGGACGATGACGAGGAGGACATCGACGACCTGGAGCACGAGTTCACATCGA 432
Db 565 GATCGTGGGAGGAGGAGACGACACTGATGCCGATAGCGACTTCAATTACCTTGCATC 624
QY 433 CGACGAGATCAGCAGAGGACGCTGGAGGGCAACATGACAGAACAGCCAGATCA----- 485
Db 625 TGGCAATGAGGACCAAGAGCAGAAAGATTGCCGACAGAAATGCGCAGTGGCGCATGAACGT 684
QY 486 -----CGAGGCGATGCTGCGCGGAGGATGAGCTACGGGAGGGGCGCCG----- 530
Db 685 TGGGGGCGAGCGGGGATGTTGGTCCGCCCAAGTATGACAGTGGCGGAGATCGGGCTTACCA 744
QY 531 --ACGACGGGACGCGCAACACACCCCGCAGATCCCGCCATCATCACCGCTCCCGCTC 588
Db 745 GTATGACAGTGGCGAGATTCTCGGGGATACATCCCATCAGTCACTACAGCCAGATCTC 804
QY 589 CGTCCCGGTGAGCGGTGAGTTTC--GATTACCAACGGGTATGGCCACGCGGAGGTCTCGT 647
Db 805 AGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGTCCCAACTGGGNAACATTGG 864
QY 648 CTTCCCTGCAACAGGCGATCCATCGGTACCTGTGTCTGAGCAGGAGGTGCAAGTGGG 707
Db 865 CAAGCGTCTCATTTTC--CCTATGTGAACCAATCCGCAAAATCCCGTCAAGGGAGTTCTCTG 923
QY 708 ACGAGAGAAAGAGTGAAGGAGGAGGATGACGACTGGAAGTCTGGAAGTCCAGCAGGCA 767
Db 924 GTAGCATTTGGGAAATGTTCCCTTGGAAAGAGAGGGTTGATGGCTTGGNAATGAGCAGGACA 983
```


Qy 768 -----TCC 770
Db 984 AGGGACGATTTCCCATGACGAATGSCAAGCAATTTGCTCTCTGAGGGTCGGGGTTTG 1043
Qy 771 TCGSCGGCGCGCGATCCGGAACACATGGAACGCGGACGTCGCACTCAAGACGAGGCGGA 830
Db 1044 GTGATATTGATGATCACTGATTACAAATGGAAGATGCTTATTGAACGACGAATCTC 1103
Qy 831 GGCAGCCGCTGTCGAGGAAGGTGCGATCGCGTGCAGCAAGGTGAACCCCGTACCGGATGG 890
Db 1104 GACAGCCTCTATCTAGGAAGTTCCACTTCCTCCAGGATAAATCCCATACAGGATGG 1163
Qy 891 TGATCGTGGTGGCTCTGTTGTCGCTCGCTCTCTCTCCGATACCGTATCTGCAACCCG 950
Db 1164 TCATTGTCGCGATGATGTTGTTCAAGCATCTCTCTGCACTACCGTATCACAAATCCTG 1223
Qy 951 TCCCGGACGCCATCGGCTGGGCTGCTCTCCATCATCTCGAGATCTGGTTGCGCATCT 1010
Db 1224 TGGCAATGATACCAATATTGGCTCTATCTGTATTATGTGAGATCTGGTTTGTCTTT 1283
Qy 1011 CTTGGATCTCGACCACTTCCCAGTGGTTCGCCATCGACCGCGAGACGTACCTCGACC 1070
Db 1284 CGTGGAATTGGATCAGTCCCTAAGTGGTTTCCAAATCAACCGGGAGACGTACCTTGATA 1343
Qy 1071 GCCTCTCCCTCAGGTACGAGGGAAGGAGGAGCGCTGCTGCTGTCGCGCGGTGGAACCTGT 1130
Db 1344 GGCCTGSCATTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTT 1403
Qy 1131 TCGTGACACGCTGAGCCGCTCAAGGAGCGCGCTGCTGACCGGCAACACCGTGTCT 1190
Db 1404 TCGTCAGTACGTGACCAATGAGGAGCTCTCTTTGTCATCTGCCAATACCGTGTAT 1463
Qy 1191 CCATCTCGCGTAGACTACCCCGTGGACAAGGTCTCTGCTAGTCTCGAGACGAGCGG 1250
Db 1464 CCATTCTGCTGGATTACCTGTGGATAAGGTCTCTTGTATGATCTGATGATGAG 1523
Qy 1251 GGTGATGCTGACGTTGAGTCTGCTGAGGAGCGCGGATTCGCGCGCAAGTGGGTGC 1310
Db 1524 CTGCGATGCTGACATTTGATGTCATAGCTAGCTGAGACTTCAGAGTTTGTGAGAAAATGGGTAC 1583
Qy 1311 CTTTCTGCAAGAGTTCCGGATCGAGCCCGCGCGCGGAGTTCTACTCTCGCTCAAGG 1370
Db 1584 CATTTGTTAGAGTACAAATTTGAACCTAGAGCTCTGATGTTGTTCTCCAGAAA 1643
Qy 1371 TCGACTACTCAAGGAAGGTGAGCCCACTTCTGTCAGGAGCGCGCGCCATGAAGA 1430
Db 1644 TTGATTACTTGAAGGAAGAGTGACCCCTCAITTTGTTAAAGACCGCGCGGCATGAAGA 1703
Qy 1431 GAGATATGAGGATTCAGGTCCGATCAACGGCTGGTGGCCAGGCGCATGAGGTGC 1490
Db 1704 GAGAATATGAAGATTCAGAGTTAGGGTAAATGGGCTTTGTTGCTAAGGCAACAGAAAGTTC 1763
Qy 1491 CGGCAAGGGGTGGATCATGAAGGACGGACCGCGTGGCGGGAACAAACCGCGAC 1550
Db 1764 CTGAGGAAGATGATCATGAGATGGGACACCATGCGCAGGAAACATACAGGGACC 1823
Qy 1551 ACCCGGCATGATCAGGTGTTCTGGGCGCACAGCGCGGCGCACGACACCGAGGGCAACG 1610
Db 1824 ATCTGGAATGATTCAGGTTTCTTGTGTCACAGTGGTGGCTTTGATCTAGTGGGCAATG 1883
Qy 1611 AGTGGCCCGCTGCTGATGCTCTGATTCGCTGTCGCGTGGAGAGCGCGGATTCAGGACCAAGA 1670
Db 1884 AGCTACCCCGTTGGTCTATGTTTCTGCTGAAAAGCGTCTCTGGATTCAGCATCACAGA 1943
Qy 1671 AGGCGGCGGCATGAACGCTGATTCGCTGTCGCGCTGCTGACCAACCGCGCATTTCA 1730
Db 1944 AAGCTGGTGCCATGAATGCTCTTGTGTTCTGATGCTGCTTACCAATGGAATPACA 2003
Qy 1731 TGCTCAACTTTGATGATCACTACATCAACACAGCAAGGCGCATCCGGAGGCGCATGT 1790
Db 2004 TGTGTAATCTTGATTGATCACTACATTAACACAGTAAAGGCTCTCAGGGAAGCTATGT 2063
Qy 1791 GCTTCTCTAGGACCCCTCAGGTGGCGGGAAGGTCTGCTAGGTTCAAGTTCCCGCAGAGGT 1850

Db 2064 GCTTCTTATGGACCCCTAACTAGGAAGGAGTGTCTGTACGTCCAGTTCCTCCAGAGAT 2123
Qy 1851 TCGACCGCATTCGACGTGACGACCGATAGCTAAACAGGAACACCGTCTCTTCGACATCA 1910
Db 2124 TCGATGGCATTCAGCAATGATCGATATGCCAACAGGAACACCGTGTCTTCGATATTA 2183
Qy 1911 ACATGAAGGGCTTGGACGCGCATCAAAGGCGCGGTGTACGTCCGGACAGGCGTGTGTTC 1970
Db 2184 ACTTGAGAGTCTTGATGGCATCCAAGGACCACTTATGTCCGAACTGGCTGTGTTTCA 2243
Qy 1971 GGGCCAGGGCTCTAGGGCTACAACCTCTCCAAAGGACCCCAAGA ----- 2015
Db 2244 ACCGAAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAGAAAGAGGGTGTCTTGT 2303
Qy 2016 -----GGCCCAAGATGCTGACCTGCGACCTGTCTGCCGTCTCTCGCGCGCAAGA 2063
Db 2304 CATCACTATGTGGCGGTAGGAGAAGCAAGCAAAATCAAAAGAGGGCTCGGACACAGAAGA 2363
Qy 2064 AGCGAAACACGCCAAGGAC -----GGGCTCCCGGAGG 2096
Db 2364 AGTCGCAAGACATGTGGACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGG 2423
Qy 2097 GCACCGCTGATATGGGATAGATAGGACACAGAGAGATGCTCATGTCCCATGAACTTCG 2156
Db 2424 GAGTTGAAGCGCTGGAATTTGACACGAGAAATFACCTTCTTATGTCTCAATGAGCCTGG 2483
Qy 2157 AGAAGCGGTTTCGGGCGAGTCCGCGCGTTCGTACGCTCGACGCTGATGGAGAAAGGCGCG 2216
Db 2484 AGAAGAGATTTGGCCAGTCCGAGCGTTTGTTCCTCCATCTGATGGAGATATGGTGGT 2543
Qy 2217 TCCCTCTTGTGTCAGACCCCGCGCTCTCTCAAGGAGGCCATCCATGTCAATCAGCTGCG 2276
Db 2544 TCCCTCAGTCCGCAACTCCGAGTCTCTCTGAAAGAGCTATCCATGTTATAAGCTGTG 2603
Qy 2277 GCTACGAGACACAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTCAATCAGCG 2336
Db 2604 GCTATGAGGACAAAGCTGAATGGGAACTGAGATCGGGTGGATCTACGGTCTGTGACAG 2663
Qy 2337 AGGACATCTGACGGGGTTCAAGATGACCTGCGCGGGTGGCGCTCGTGTACTGATGC 2396
Db 2664 AAGACATTTCTACCGGATTCAGATGCAACGGGAGGCTGGCGGTGATCTACTGATGC 2723
Qy 2397 CGAAGCGGCGGCTTCAAGGGTTCGCGCGCATCAATCTATCGGACCGCTCTCAACGAG 2456
Db 2724 CCAAGCGCGCAGCTTCAAGGGGTCTGCCCATCAATCTTTCGACCGCTCTGAACCCAG 2783
Qy 2457 TGTCTCGGTGGGCTGGGGTTCGTCGAGATCTTCTTTCAGCGCGGACAGCCCCCTGCTGT 2516
Db 2784 TGTCTCGGTGGGCTCTTGGGTCCGTGGAGATCTCTTTCAGCGCGGCACTGCCCCCTGTGGT 2843
Qy 2517 ACGGCTTACAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCGCTACATCAACACCA 2576
Db 2844 ACGGCTAC ---GGAGGGCGGCTCAAGTCTCTGGAGAGATTCGCGTATCAACACCA 2900
Qy 2577 TCTACCCCTTCACTCGCTCCGCTGCTGCTACTGCAACCTCTCCCGCGCTCTGCTCC 2636
Db 2901 TCTACCGCTCACTCCGCTTCTCATCTACTGCACTCTCTGCGCGCATCTGTCTGC 2960
Qy 2637 TCACCGCAAGTTTCATCATGCTGCTGATAGACGTTTCGCGACGCTCTTCTTTCATCGCCC 2696
Db 2961 TCACCGGAAAGTTTCATCTCCAGAGATCAGCAACCTTCGCGCAGCATCTGGTTCATCTCCC 3020
Qy 2697 TCTTCATGTCATCTTTCGCGACGGGATCTCTGGAGATTCGGGTGGAGCGGGTGGAGCATCG 2756
Db 3021 TCTTCATCTCGATCTTCGCCACGGGATCTCTGGAGATGAGTGGAGCGGGTGGGATCG 3080
Qy 2757 AGGAGTGGTGGAGAACGAGCAGTTCGTGGGTCAATCGCGCGCGTGTCTCGGCACTCTCTTCG 2816
Db 3081 ACGAGTGGTGGAGAACGAGCAGTTCGTGGGTGATCGGGGCGCATCTCGCGCAGCTCTCTTCG 3140
Qy 2817 CGGTCTGACGGGCTGCTCAGGTCTCGCGGGATCGACACCAACTTCACCGTCACTCT 2876


```
Db 3141 CCGTGTTCACGGGCTGCTCAAGTGTCTGGCGGCGATCGACACCACTTACCGTCACCT 3200
Qy 2877 CCAAGGCCACCGGCGACGAGGACGACGAGTTTCGGCGAGCTCTACGCCCTTCAAGTGGACCA 2936
Db 3201 CCAAGGCCCTCG---GACGAGGACGGCGACTTTCGGCGAGCTGTACATGTTCAAGTGGACGA 3257
Qy 2937 CGCTCCTCATCCCGCCACACGCGTGTCTCATCATTAAGTTCATCGGCGTGTGGCGCGCA 2996
Db 3258 CGCTCCTGATCCCGCCACACCACTCCTGATCATCAACCTGGTCGGGCGTGTGGCGCGCA 3317
Qy 2997 TCTCCGAGCCCATCAACACGGGTACCAAGTCTCTGGGGGCCCTCTTCGGCAAGCTCTTCT 3056
Db 3318 TCTCCTAGCCATCAACAGCGGATACCAAGTCTGTGGGGCCCGCTCTTCGGCAAGCTCTTCT 3377
Qy 3057 TCGCCTTCTGGGTCAATCGTCCACCTCTACCCGTTCTCAAGGGGCTCATGGGGCGCCAGA 3116
Db 3378 TCGCCTTCTGGGTCAATCGTCCACCTGTACCCGTTCTCAAGGGGCTCATGGGGCAGGAGA 3437
Qy 3117 ACAGGAGCCCGACCGTTGTTGTCATCTGCTCCATCTGCTGGCGCTCCATCTTCTCCCTGC 3176
Db 3438 ACCGACCCCGACCATCGTGTGCTGTGGGCCATCCTGCTGGCGTCCATCTTCTCCCTGC 3497
Qy 3177 TCTGGGTGAGGATCGACCCCTTTCATCGTCAAGGACCAAGGGGCCCGGACGTCAAGGAGGTG 3236
Db 3498 TGTGGGTTCGATCGACCCCTTACACCCCGGTCACTGGGCCCGGATACCCAGAGGTG 3557
Qy 3237 GCATCAATTGCT 3248
Db 3558 GCATCAACTGCT 3569
```

Search completed: March 12, 2005, 02:52:34
Job time : 1256 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 15:50:25 ; Search time 7012 Seconds
(without alignments)
18690.143 Million cell updates/sec

Title: US-10-627-132-29
Perfect score: 3443
Sequence: 1 ctgcgtgcctgcctgcga.....aaaaaaaaaaaaaaaaaaaaa 3443

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_hrc.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gss1.*

9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2689.8	78.1	3168	9	CL976367
2	1398.8	40.6	3763	3	AY108113
3	1347.6	39.1	3788	3	AY103701
4	1272.6	37.0	3192	9	CL972423
5	1266.4	36.8	3783	3	AY104730
6	1259.4	36.6	3728	3	AY112236
7	1201.4	34.9	3897	3	AY103655
8	1197.2	34.8	3898	3	AY110415
9	1132.8	32.9	3696	3	AY110079
10	1045.2	30.4	3911	3	CNSO1ML
11	1033	30.0	2872	3	AY104236
12	791.2	23.0	2595	9	CL964957
13	782.4	22.7	1874	6	CD726831
14	757	22.0	789	7	CF430024
15	747.6	21.7	943	9	CG454544
16	707.6	20.6	1189	3	AY107656
17	695.8	20.2	872	8	CC427226
18	680.4	19.8	834	7	CO519653
19	679.2	19.7	710	6	CA261021
20	679.2	19.7	3474	9	CL978864
21	670.8	19.5	949	9	CG211396
22	646.4	18.8	822	9	CG198455
23	645.2	18.7	687	6	CA127439
24	644.6	18.7	706	6	CA153363

25	633	18.4	781	6	CA129877
26	632.8	18.4	715	6	CA069753
27	631.4	18.3	3048	9	CL974686
28	631	18.3	772	6	CB646549
29	623.2	18.1	693	6	CA235459
30	622.2	18.1	746	6	CA155520
31	614.6	17.9	821	9	CG383072
32	607.4	17.6	777	9	CL167974
33	603.4	17.5	631	6	CA133799
34	599.4	17.4	982	9	CC615547
35	599	17.4	643	6	CA244716
36	597.6	17.4	935	9	CC722858
37	588	17.1	758	6	CA183524
38	587.2	17.1	791	6	CB661447
39	587	17.0	845	9	CG215949
40	582.4	16.9	749	6	CA157562
41	582.2	16.9	619	6	CA211120
42	581.4	16.9	945	9	CG300346
43	580.4	16.9	790	6	CB643324
44	577.8	16.8	911	8	CC350858
45	576.2	16.7	943	9	CC660675

ALIGNMENTS

RESULT 1

CL976367

LOCUS

CL976367

DEFINITION

OsIFCC028793 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.

ACCESSION

CL976367

VERSION

CL976367.1

KEYWORDS

GSS.

SOURCE

Oryza sativa (indica cultivar-group)

ORGANISM

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 3168)

AUTHORS

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.

TITLE

An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis

JOURNAL

Unpublished (2004)

COMMENT

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

source

1..3168

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Express Library"

/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match

Best Local Similarity

Matches 2899; Conservative

78.1%; Score 2689,8; DB 9; Length 3168;

91.4%; Pred.No. 0;

Mismatches 257; Indels 15; Gaps 4;

Qy

92 ATGAGGCCACCGCCGGCGTGGTGGCGGCTCGCACACCGAACGAGCTGGTCTGATC 151

Db

1 ATGAGGCCACCGCCGGCGTGGTGGCGGCTCGCACACCGAACGAGCTGGTCTGATC 60

Qy

152 CGGGGGCCACGAGGACCCCAAGCCGCTCGCGGGCGCTGAGCGGGCAGGTGTGCGAGATATGC 211

1118 GTGTCGTCTACGTCTCCGACGACGGCGCTCCATGCTCACGTTCCGAGTCGCTGTCGGAG 1197
1283 AGGCGCGAGTTCCGCGCAAGTGGGTGCCCTTCTGCAAGAGTTCCGCAATCGAGCCCCGC 1342
1198 AGCGGGGAGTTTCGCGCGAAGTGGGTGCCATTTCTGCAAGAGTTTCAGCATCGAGCCCCGC 1257
1343 GCCCGGGAGTTTACTTTCTTCGCTCAAGGTTCGACTACTCTCAAGGACAAGGTTCATCCCAAC 1402
1258 GCCCGGGAGTTTACTTTCTTCGCTCCAGAGGTTCGACTACTCTCAAGGACAAGGTTCATCCCAAC 1317
1403 TTCGTGACGAGGCGCGCCCATGAAGAGAGATGAGAGAGTTCAAGTTCGGATCAAC 1462
1318 TTCGTGACGAGGCGCGCCCATGAAGAGAGATGAGAGAGTTCAAGTTCGGATCAAC 1377
1463 GCGCTGTCGCGCAAGGCCATGAAGGTTCGCGAGAGGGGTGATCATGAAGGACGGCAGC 1522
1378 GCGCTGTCGCGAGGGCGAGAGGTTCGCGAGAGGGGTGATCATGAAGGACGGGAGC 1437
1523 CCGTGGCCCGGGAAACAACACCCCGGACCAACCCCGGCGATGATCCAGGTGTTCTTGGGCGAC 1582
1438 CCATGGCCCGGGAAACAACACCCCGGACCAACCCCGGCGATGATCCAGGTGTTCTTCGGCCAC 1497
1583 AGGCGGCGCGACGACCGAGGGCAACCGAGCTGCCCGGCTCGTGTACGTTCTCCCGTGAG 1642
1498 AGCGGCGGCGACGACACCGAGGGCAACCGAGCTGCCCGGCTCGTGTACGTTCTCCCGTGAG 1557
1643 AAGCGCCCGGGATTTCCAGCACCAAGAGAGGGCGCGCATGAACGCTCTGTATTCGCGTC 1702
1558 AAGCGCCCGGCTTCGACCAACAAGAGGGCGCGCATGAACGCTCTGTATTCGCGTC 1617
1703 TCCGCGCTGTCGACCAACCGCGCTTCTGCTCAACTTGGACTGTGATCACTACATCAAC 1762
1618 TCGCGCGCTGTCGACCAACCGCGCTTCTGCTCAACTTGGACTGTGATCACTACATCAAC 1677
1763 AACAGCAAGGCCATCCGAGAGGCGCATGTGTTCTCATGAGCCCTCAGTTCGCGCGGAAG 1822
1678 AACAGCAAGGCCATCCGAGAGGCGCATGTGTTCTCATGAGCCCGCAGGTTCGACCGAAG 1737
1823 GTCTGTAGTTTCAGTTCCGCGAGGTTTCGAGGATCGAGGATCGAGGATCGACCATACGCT 1882
1738 GTGTGCTAGTTTCGAGTTCCGCGAGGTTTCGAGGATCGAGGATCGAGGATCGACCATACGCT 1797
1883 AACAGCAACACCGCTTCTTTCGACATCAACATGAAGGGGTTCGACCGCATCCAAAGGCCG 1942
1798 AACCGCAACACCGCTTCTTTCGACATCAACATGAAGGGGTTCGATGGATCCAGGGCGCG 1857
1943 GTGTAGTTCGCGACAGGTTCGTTCCGCGCGAGGCGCTCTACGCTACACCCCTCC 2002
1858 GTGTAGTTCGCGACAGGTTTCGAGGATCGAGGATCGAGGATCGACCATACGCT 1917
2003 AAGGGACCCAAAGAGGCCCAAGATGTTGACCTGCGACTGCTGCGCGCTTCGCGCGCAAG 2062
1918 AAGGGACCCAAAGAGGCCCAAGATGTTGACCTGCGACTGCTGCGCGCTTCGCGCGCAAG 1977
2063 AAGCGGAAACACCGCCAAAGGACGGGCTCCCGAGGCGCTCCCGCGCGAGCGGAGTGGAC 2119
1978 AAGCGGAAACACCGCCAAAGGACGGGCTCCCGAGGCGCTCCCGCGCGAGCGGAGTGGAC 2037
2120 AGCGCAAGAGATGCTCATGTCCCATGAACTTCGAGAGGCTTCGGGCGAGTCCGCG 2179
2038 AGCGCAAGAGATGCTCATGTCCCATGAACTTCGAGAGGCTTCGGGCGAGTCCGCG 2097
2180 GCGTTGCTGCTGCTGCTGATGAGAGAGGGCGGCTCCCTCTCTGCTCGAGCCCGCG 2239
2098 GCGTTGCTGCTGCTGCTGATGAGAGAGGGCGGCTCCCGCGCTGCTCCAGCGCGCG 2157
2240 GCGCTCTTCAAGAGGCGCATTCATGTCTCAGCTGCGGCTTACGAGGACAAGACCGACTGG 2299
2158 GCGCTCTTCAAGAGGCGCATTCATGTCTCAGCTGCGGCTTACGAGGACAAGACCGACTGG 2217
2300 GCGCTGAGTGGGTGGATCTACGGGTTCATCGAGGACATCTTGAAGGGGTTCAAG 2359
2218 GGTCTCGAGTGGGTGGATCTTACGGGTTCATCGAGGACATCTTAAAGGGGTTCAAG 2277

61 CGGGGCAAGAGGAGCCCAAGCGCTCGGGGCTGAGCGGCGAGGTGTCGAGATATGC 120
212 GCGGACGAGGTTCGGGCTCAGCGTGAACGGGCACTTTCGTGCTGCTGCAACGAGTCCGGC 271
121 GCGGACGAGGTTCGGGCTCAGCGTGAACGGGCACTTTCGTGCTGCTGCAACGAGTCCGGC 180
272 TTCCTGCTGTCGGGCTTCGCTACGAGTACGAGCGCGCGGAGGCGACGAGAACTCCCG 331
181 TTCCTGCTGTCGGGCTTCGCTACGAGTACGAGCGCGCGGAGGCGACGAGAACTCCCG 240
332 CAGTGAAGACGCGCTTCAAGCGCTTCAAGGGGAGCCCGAGGGTTCGCGGGACGATGAC 391
241 CAGTGAAGACGCGCTTCAAGCGCTTCAAGGGGAGCCCGAGGGTTCGCGGGAGCAGGAC 300
392 GAGGAGGACATCGACGACCTGAGGACGAGTTTCAATCATCGACGACGAGATCAGCAGAG 451
301 GAGGAGGACATTCGACGACCTGAGGACGAGTTTCAATCATCGACGACGAGGAGCAGAG 360
452 CAGCTGAGGGCAA-----CATGCAGAACAGCCAGATCAACGAGGCGATGCTGCAACGGC 505
361 CTGACAGGATCAGGATGCGATGCGAGACGAGCAATCAACGAGGCGATGCTGCAACGGC 420
506 AGGATGAGTACCGGAGGGGCCCCGACGAGCGGCGACGCAACCAACCCCGCAGATCCCG 565
421 AAGATGAGTACCGGAGGGGCCCCGACGAGCGGCGACGCGCAACAGCACCCCGC---TCCCG 477
566 CCATCATACCCGGCTCCCGCTCCGTCGCGGTGAGCGGTGAGTTTCGATTAACAACGG 625
478 CCGATCATACCCGGCTCCGCTCCGTCGCGGTGAGCGGGAGTTTCGATTAACAACAGC 537
626 TATGGCCACGGCAGGTCTGCTTCCCTGCAAGCGCATCAATCCGTACCCGTGTCT 685
538 CATGGCCATGGCAGTTCTCTCTTCTGCAAGCGCATCAACCCCTACCCCGTGTCT 597
686 GAGCCAGGAGTCCCAAGTGGGACGAGAAAGAAAGTGAAGTTCGGAAGGAGAGGATGGAC 745
598 GAGCCAGGAGTGCMAAGTGGGACGAGAGAAAGAGTGAAGTTCGGAAGGAGAGGATGGAC 657
746 GACTGGAAGTCCAAAGCGAGGCGATCTTCGCGCGCGGGG---CCGATCCGGAAGACATGGAC 802
658 GACTGGAAGTCCAAAGCGAGGCGATCTTCGCGCGCGGGGCGCCCGATCCCGACGACTACGAC 717
803 GCGGAGTGGCATGCAACGAGAGGCGAGGCGCGCTGTCGAGGAGGTTTCGATCGCG 862
718 GCGGAGTGGCATGCAACGAGAGGCGAGGCGCGCTGTCGAGGAGGTTTCGATCGCG 777
863 TCAGGCAAGGTGAACCCGTTACCGGATGGTGAATCGTGGTGGCTCTCGTTGCTCGCTTC 922
778 TCAGGCAAGGTGAACCCGTTACCGGATGGTGAATCATCTCTCGTCTCGTGGTGGCTCGGCTTC 837
923 TTCCTCGGTACCGGTATCTGCAACCGCTCCCGGACCGCATTCGGGCTGTGGCTCGTCTCC 982
838 TTCCTCGGTACCGGTATCTGCAACCGCTCCCGGACCGCATTCGGGCTGTGGCTCACTCC 897
983 ATCATCTGCGAGATCTGTTTCGCTTCGCTGATCTTCGACAGTTCGCCAAGTGGTTC 1042
898 ATCATCTGCGAGATCTGTTTCGCTTCGCTGATCTTCGACAGTTCGCCAAGTGGTTC 957
1043 CCCATCGACCGCGAGATCTCCTCGACCGCTCTCCCTCAGGTACGAGAGGGAAGGGGAG 1102
958 CCGATCGACAGGAGACCTACCTCGACCGCTCTCCCTCGCTACGAGCGGAGGGGAG 1017
1103 CCGTCTGCTGTTCGGGCTGGACCTGTTTCGTGAGCAGCGGTGAGCCCGCTCAAGGAGCCG 1162
1018 CCGTCTGCTGTTCGGGCTGGACCTGTTTCGTGAGCAGCGGTGAGTCCCGCTCAAGGAGCCG 1077
1163 CCGTCTGCTGACCGCAACCGCTCTCTCATCTCGCGGTGAGTACCTACCCGCTGAGCAAG 1222
1078 CCGTCTGCTGAGGCAACCGCTCTCTCATCTCGCGGTGAGTACCTACCCGCTGAGCAAG 1137
1223 GTCTCTGCTACGTTCCGACGAGCGGCGCTCGATGCTGAGCTTCGAGTCTGCTCGGAG 1282

```
Qy 2360 ATGCACTCCGCGGCTCGGCTCGTGTACTGATGCGAAGCGGCGCGCTTCAAGGGG 2419
Db 2278 ATGCACTCCGCGGCTGAGATCGTGTACTGATGCGAAGCGGCGCGCTTCAAGGGG 2337
Qy 2420 TCGGCGCGGATCAATCTATCGGACCGTCTCAACAGAGTGCTCCGGTGGCGCTGGGTCC 2479
Db 2338 TCGGCGCGGATCAACCTATCTGACCGTCTCAACAGAGTGCTCCGGTGGCGCTCGGTCC 2397
Qy 2480 GTCGAGATCTTTTACGCGGCACAGCCCTCTGTGTACGGCTCAAGAAGCGCAACCTC 2539
Db 2398 GTCGAGATCTTTTACGCGGCACAGCCCTCTGTGTACGGCTCAAGAAGCGCAACCTT 2457
Qy 2540 AAGTGGCTGGAGCGCTTCCCTACATCAACACCATCTACCCCTTCACTCGTCCCG 2599
Db 2458 AAGTGGCTGGAGCGCTTCCCTACATCAACACCATCTACCCCTTCACTCGTCCCG 2517
Qy 2600 CTGCTCGGCTACTGACACCTCCCGCGCTCTGCTCTCCCTCAACCGGCAAGTTCATGCGG 2659
Db 2518 CTCTCGGCTACTGACACCTCCCGCGCTCTGCTCTCCCTCAACCGGCAAGTTCATGCGG 2577
Qy 2660 TCGATTAGCACGTTGCGCAGCTCTTCTTCATCGCCCTCTTTCATCTCCATCTTCGCGACG 2719
Db 2578 CCGATTAGCACGTTTGGAGTTTGTCTTCATCGCGCTTTCATCTCCATCTTCGCGACG 2637
Qy 2720 GGCATCTGGAGATCGGCTGAGCGGGTGAAGCATCGAGGAGTGGTGGAGGAAAGAGCAG 2779
Db 2638 GGCATCTGGAGATGAGTGGAGCGGGTGAAGCATCGAGGAGTGGTGGAGGAAAGAGCAG 2697
Qy 2780 TTCTGGGTCAATCGCGCGCTCGCGCATCTTTCGCGCTCTTCGCGCTGTCAGGGCTCTCAAG 2839
Db 2698 TTCTGGGTCAATCGCGCGCTCGCGCATCTTTCGCGCTCTTCGCGCTGTCAGGGCTCTCAAG 2757
Qy 2840 GTCTCGCGGGATCGACCAACTTCAACCTCTCACTCCAGGCGCACCGGCGAGGAGAC 2899
Db 2758 GTGCTGGCGGGATCGACCAACTTCAACCTCTCACTCCAGGCGCACCGGCGAGGAGAC 2817
Qy 2900 GACGAGTTGCGCGAGCTTACGCTTCAAGTGAACAGCTCTTCATCCCGCGCCACCAAG 2959
Db 2818 GACGAGTTGCGCGAGCTTACGCTTCAAGTGAACAGCTCTTCATCCCGCGCCACCAAG 2877
Qy 2960 CTGCTCATATTACGTCATCGGCGTGTGGCGGCTCTTCGCGCTCTTCGCGCTTTCGCGCT 3019
Db 2878 CTGCTCATCTTCAACATCATCGGCGTGTGGCGGCTCTTCGCGCTTTCGCGCTTTCGCGCT 2937
Qy 3020 TACCAGTCTCGGCGCGCTCTTCGCGCAAGCTCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 3079
Db 2938 TCCGAGGCTGGGCGCGCTCTTCGCGCAAGCTCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 2997
Qy 3080 CTCTACCGGTTCTCAAGGGGCTCATGGGCGCGCAGAACAGGACCGCCACCGGTTGTGTC 3139
Db 2998 CTCTACCGGTTCTCAAGGGGCTCATGGGCGCGCAGAACAGGACCGCCACCGGTTGTGTC 3057
Qy 3140 ATCTGCTCCATCTGCTGGCTCCATCTTCTCCCTGCTCTGGGTGAGGATGACCTTTC 3199
Db 3058 ATCTGCTCCGCTGCTGCTGCTCCATCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3117
Qy 3200 ATCTGCTGAGGACCAAGGGGCGCGGCTCAGGAGTGTGGCATCAATTTGCTGA 3250
Db 3118 ACCATCAAGGCGCGGCGCTGAGTCAAGGAGTGGGCGATCACTGCTGA 3168

RESULT 2
AY108113
LOCUS AY108113 3763 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays P0126465 mRNA sequence.
ACCESSION AY108113
VERSION AY108113.1 GI:21211191
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCA
```

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 3763)

Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,

Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of

Overgo Probes

Unpublished (2002)

2 (bases 1 to 3763)

Coe, E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones,

these are publicly available from ZmDB and may be found by BLAST

searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat

Schnable, Iowa State, then clones may be requested from ZmDB:

www.zmdb.iastate.edu.

Location/Qualifiers

FEATURES

source

1. 3763

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="MaizeDB:637367"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/DuPont Consensus

Library"

/notes="this sequence is part of a project of EST

assemblies resulting from the application of public

contigs to seed DuPont contigs; this resource was

assembled by DuPont as part of a collaboration for the

overgo addressing of BACs in conjunction with the Maize

Mapping Project"

ORIGIN

Query Match 40.6%; Score 1398.8; DB 3; Length 3763;

Best Local Similarity 67.2%; Pred. No. 2.9e-266;

Matches 2185; Conservative 1; Mismatches 923; Indels 143; Gaps 9;

Qy 133 GAACGAGCTGCTGCTGATCGGGGCCACGAGGACCCCAAGCGCTCGGGCGCTGAGCGG 192

Db 325 GAGCTCGCTTCCATGAGGGGCGACGCGGAGCGGTGAGTCGGGGAGGCGCGTGGCGG 384

Qy 193 GCAGGTGTGCGAGATATGCGCGCAGCAGAGTTCGGGCTCACGTGCGACCGCAGCTTTCGT 252

Db 385 ACAGGTGTGCGAGATCTGCGCGCAGCGCTGGGCACCAACGCGGAGGAGCTTTCGC 444

Qy 253 CGCTGCAACAGTGGCGCTTCCCGCTGTCGGCCCTGCTACGAGTACGAGCGCGCGGA 312

Db 445 CGCTTGCAGCTCTGCGGGTTCGGGTGTCGGCCCTGCTACGAGTACGAGCGCAAGGA 504

Qy 313 GGGCAGCAGAACTGCCCCAGTGCAGACGCGCTACAAGCGCTCAAGGGAGCCCGAG 372

Db 505 CGCAGCAGCGGCTGCCCCAGTGCAGACCAAGTACAGGCGCACAGGGAGCCCGC 564

Qy 373 GGTTCGGGGAAGATGACGAGGAGGACATCGACGCTCGAGCAGCAGGTTCAACATCGA 432

Db 565 GATCCGTGGGGAAGGAGACGACACTGATGCGGATGCGGCTTCAATACCTTGCATC 624

Qy 433 CGAGGAGATCAGCAGAGGAGCTGGAGGGCAACATGCAACAGCCAGATCA----- 485

Db 625 TGGCAATGAGGACCAAGAGCAAGATTGCGGACAGAAATCGCAGCTGGCGCATGAACGT 684

Qy 486 -----CCGAGGGGATGCTGTCACGGCAGGATGAGCTACGAGGGGCGCCCG----- 530

Db 685 TGGGGGAGCGGGGATGTTGTCGCCCAAGTATGACAGTGGCGAGATCGGGTTACAA 744

Qy 531 --ACGAGCGGAGCGCAACCAACCCCGCAGATCCCGCCCATCATCACCGGCTCCCGCTC 588

Db 745 GTATGACAGTGGGAGATTCTTCGGGATACATCCCATCATCAGTCACTAACGAGCATCTC 804

Qy 589 CTGCGCGGTGAGCGGTGAGTTTCC-GATTACCAACGGGTATGGCCACGCGAGTCTCGT 647

Db 805 AGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGTCCCACTGGGAACATTGG 864
Qy 648 CTTCCCTGCACAGCGCATCCATCCGTACCTGTGTCTGAGCCAGGAGTGCACAGTGG 707
Db 865 CAAGCGTGTCCATTTTC - CCTATGTGAACCAITTCGCCAAATCCGTCAAGGAGTTCTCTG 923
Qy 708 ACGAGAGAAGAAAGTGCAGTGAAGGAGAGGATGGACGACTGGAAATGTCACAGAGGCA 767
Db 924 GTAGCATTTGGGATGTTGCCCTGGAAGAGAGGTTGATGGCTGGAAATGACACAGACA 983
Qy 768 -----TCC 770
Db 984 AGGGAGCAATCCCATGACGAATGGCACAGCATTGTCTCCCTCTGAGGGTCCGGGTGTTG 1043
Qy 771 TCGGGGGGGCGGATCCGAGAGACATGGAACCGGACGTGGCACTGMAAGCAGGCGGA 830
Db 1044 GTGATATTTGATGATCAACTGATTACAAACATGGAAGATGCTTATTGAACGACGAACCT 1103
Qy 831 GGCAGCGCTGTGCGAGGAAGTGTGATCGGTGCGAGCAAGTGAACCCCGTACCAGATGG 890
Db 1104 GACAGCTCTATCTAGGAAGTTCCACTTCTTCTCCAGGATTAATCCATACAGATGG 1163
Qy 891 TGATCGTGTGCTGTGCTGTGCTGCTGCTTCTTCTCCGTTACCGGTATCCTGCAACCCG 950
Db 1164 TCATTGTGCTGCGATTGATTGTTCTAAGCATCTTCTTGCACTACCGGTATCACAAATCCTG 1223
Qy 951 TCCGGAGCCATCGGGCTGTGCTGTCTCCATCATCTGCGAGATCTGGTTCGCCATCT 1010
Db 1224 TGGCAATGCATACCCATTAATGCTTCTATCTGTATTATGTGAGATCTGGTTTGTCTTT 1283
Qy 1011 CTTGGATCTCTGACCAAGTTCCCAAGTGGTTCCCACTGACCGCGAGACGTACTCTGACC 1070
Db 1284 CGTGGATATTGGATCAGTTCCCTTAAGTGTTCCTATCAACCGGAGACGTACTTTGATA 1343
Qy 1071 GCCTCTCCCTCAGGTACGAGAGGAAGGAGCGGTGCTGTGCTGTCGGCGGTGACCTGT 1130
Db 1344 GGCTGGCATTAAGTATGACCGGGAAGGTGAGGCATCTCAGTTTGGCTGTGTGACATTT 1403
Qy 1131 TCGTGACGCTGGACCCGCTCAAGAGCGCGCGTGTGACCGCCACACACGCTCTCT 1190
Db 1404 TCGTCACTAGTGCACCCAAATGAAGAGCTCTCTTGTCACTGCCCAATACCGTGTCTAT 1463
Qy 1191 CCATCTCTCGCGTAGACTACCCGCTGGAACAAGTCTCTCTGCTAGCTCTCCGAGCAGCGG 1250
Db 1464 CCATCTTGTGTGGATTACCTGTGTAAGGTCTCTCTGCTATGATCTGATGATGGAG 1523
Qy 1251 CGTCAATGCTGACGTTTCAGTGTGCTGTGCGAGACGGCCGAGTTCGCGCGCAAGTGGTGC 1310
Db 1524 CTGCAATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTTGTAGAAAATGGGTAC 1583
Qy 1311 CCTTCTGCAAGAGTTTCGGCATCGAGCCCGCGCCCGGAGTTCTACTTCTCGTCAAGG 1370
Db 1584 CATTTGTGAAGAGTACAAATTTGAACCTAGAGCTCTCTGAATGGTACTTCTCCAGAAAA 1643
Qy 1371 TCGACTACTCTCAAGGACAAGTGCAGCCCACTTCGTGCGAGGAGCGCGCGCCATGAAGA 1430
Db 1644 TTGATTAATTGAGGACAAGAGTGCACCTTCATTTGTTAAGACCGCGCGGCCATGAAGA 1703
Qy 1431 GAGAGTATGAGAGTTCAAGGTCCGATCAACCGCTGGTGGGCAAGCCATGAAGTGC 1490
Db 1704 GAGAAATGAAGAAATCAAGTTAGGGTAAATGGCTTGTGCTTAAGGCAACAGAAATTC 1763
Qy 1491 GGCAGAGGGGTGATCATGAGACGGCAGCGCGTGGCCCGGGAACAACACCCGCAAC 1550
Db 1764 CTGAGGAAGGATGGATCATGCAAGATGGCACACCATGGCCAGGAAAAAATACMGGGACC 1823
Qy 1551 ACCCGGCATGATCCAGGTGTTCTTGGGCCACACGCGGGGCCAGACACCGAGGGCAAG 1610
Db 1824 ATCTGGAATGATTACAGTTTCTTGGTACAGTGGTGGCTTGTACTGAGGGCAATG 1883
Qy 1611 AGTCCCGCGCTCGTGTAGCTTCCCGTGAAGAGCGCCGGGATTCAGCACCAACAAGA 1670
Db 1884 AGCTACCCCGTTTGGTCTATGTTTCTGTGAAAAGCGTCTCTGGATTCCAGCATCACAGA 1943

Qy 1671 AGGCGGGCCCATGAACGCTCTGATTTGGGTCTTCGCCGTGTGTGACCAACGCGCCATTCA 1730
Db 1944 AAGCTGGTGCATGATGCTCTTGTTCGTGTCTCAGCTGTGCTTACCAATGGACAATACA 2003
Qy 1731 TGTCTCAATTTGAGTGTGATCACTACATCAACAGCAAGGCCATCCGGAGGCCATGT 1790
Db 2004 TGTGGAATCTTGATTTGATCACTACATTAACAAAGTAAGGCTCTCAGGGAAGCTATGT 2063
Qy 1791 GCTTCTCATGAGACCTCAGGTTCGGCCGAAGTCTGTCTACGTTTCAGTTTCCCGCAGAGT 1850
Db 2064 GCTTCTTATGAGACCTAACCTAGGAAGGAGTGTCTGTACGTCCAGTTTCCCGCAGAT 2123
Qy 1851 TCGACGGCATCGAGGTGACGACCGGATAGCTTAAACAGGAACACCGTCTTCTTCGACATCA 1910
Db 2124 TCGATGGCATTTGACAGGAATGATCGATATGCCACAGGAACACCGTGTTCGATATTA 2183
Qy 1911 ACATGAAGGGGTGGACGATCCAAAGGCCGGTGTACGTTCGGAACAGGGTGTGTTC 1970
Db 2184 ACTTGAGAGGTCTTGATGGCATCCAGGACCAAGTTTATGTTCGAACTGGCTGTGTTC 2243
Qy 1971 GGGCCAGGGCTCTACGGCTACACCCCTCCCAAGGGACCCAGA----- 2015
Db 2244 ACCGAAACGCTCTATATGTTATGAGCCCCCAATTAAAGCAGAAAGGGTGTGTCTTGT 2303
Qy 2016 -----GGCCCAAGATGGTGACCTGCGACTGTCTGCCGTGTCTCGGCCGCAAGA 2063
Db 2304 CATCATATGTGGCGGTAGGAGAGGCAAGCAAAATCAAAGAGGGCTCGGACAAGAGA 2363
Qy 2064 AGCGGAAAACGCCAAAGGAC-----GGGCTCCCGGAGG 2096
Db 2364 AGTCGAGAAAGCATGTGGAACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGAGG 2423
Qy 2097 GCACCCCTGATATGGGAGTAGTAGGACAGAGAGATGCTCATGTCCCAATCAACTTCG 2156
Db 2424 GAGTTGAAGCGCTGGAATTTGACGACGAGAAATCACTTCTTATGTCTCAAAATGAGCCTG 2483
Qy 2157 AGAAGCGGTTTCGGGCGAGTCCGCGGCTTGTTCAGTTCGACGTGTGATGGAGAAAGGGCGG 2216
Db 2484 AGNAGAGATTTGGCAGTCCGACAGCTTGTTCCTTCCCTCCTCTGATGGAGTATGTTG 2543
Qy 2217 TCCTCTCTTGTGAGCGCCCGCGCTCCTCAAGGAGGCATTCATGTCTCAGCTGCG 2276
Db 2544 TTCTCAGTCCGCAACTCCGGAGTCTCTTCGAAAGAAAGTATCCATGTTTATAAGCTGTG 2603
Qy 2277 GCTACGAGGACAGACCGACTGGGGCTGGAGCTGGGGTGGATCTACGGGTTCGATCACGG 2336
Db 2604 GCTATGAGGACAAAGCTGAATGGGGAACCTGAGATCGGGTGGATCTACGGTTCTGTGACAG 2663
Qy 2337 AGGACATCTGACGGGTTCAAGATGCATTCGCCCGGGTGGCGCTCCGTGTACTGCAATGC 2396
Db 2664 AAGACATTTCTACCGGATTCAGATGCAACGCGGAGGCTGGCGGTGATCTACTGCAATGC 2723
Qy 2397 CGAAGCGGGCGGTTCAAGGGGTTCGGCGCGGATCAATCTATTCGACCGCTCTCAACGAG 2456
Db 2724 CCAAGCGCGCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTTCGGAACGCTCTGAACGAG 2783
Qy 2457 TGTCTCGGTGGGCGCTGGGGTTCGTCGAGATCTTCTTCAGCGCGGCAAGCCCTGCTGT 2516
Db 2784 TGTCTCGGTGGGCTCTTGGGTCCGTGGAGATCTCTTTCAGCCGGCATCTGCCCTCTGTGGT 2843
Qy 2517 ACCGCTACAAAGAACGGCAACTCAAGTGGCTGGAGGCTTCGCGCTCATCAACACCAACCA 2576
Db 2844 ACCGCTAC---GGAGGGCGGCTCAAGTTCTGGAGAGATTCGGGTACATCAACACCAACCA 2900
Qy 2577 TCTACCCCTTCACTCGCTCCCGTGTCTGCTCTACCTGTGCAACCTCCCGCGCTCTGCTCC 2636
Db 2901 TCTACCGCTCAGGTCCATCCCGCTTCTCATCTACTGTCATCTCTGCGCCGCTCTGTCTGC 2960
Qy 2637 TCAACCGCAGTTTCATGCTCGGATTAGCAGTTTCGCGAGGCTCTTCTTCAATCGCC 2696
Db 2961 TCACCGGAAAGTTTCATCTTCAGAGATCAGCAACTTCGCGAGCATCTGGTTCTATCTCCC 3020


```
QY 2697 TCTTCATGTCATCTTCGCGACGGGCATCTGGAGATGGGTGGAGCGGGGTGAGCATCG 2756
Db |||||
QY 3021 TCTTCATCTCGATCTTCGCGACGGGCATCTTGAGATGAGGTGGAGCGGGGTGGGCATCG 3080
Db |||||
QY 2757 AGGAGTGTGGAGGAACAGACAGTCTTGGGTCTATCGGGCGGTGTCGGCGCATCTCTTCG 2816
Db |||||
QY 3081 ACGAGTGTGGAGGAACAGACAGTCTTGGGTGATCGGGGCATCTTCGGGCACCTCTTCG 3140
Db |||||
QY 2817 CCGTGTGGAGGGCTGTCAAGGTCTCGCGGGATCGACCAACTTCAACCGTCACCT 2876
Db |||||
QY 3141 CCGTGTTCAGGGCTGTCAAGGTCTCGCGGGATCGACCAACTTCAACCGTCACCT 3200
Db |||||
QY 2877 CCAAGGCCACCGGCACGAGGACGAGTTCGCGGAGTCTACCGCTTCAAGTGACCA 2936
Db |||||
QY 3201 CCAAGGCTCG---GACGAGGACGGGAGCTTCGCGGAGCTGTACATGTTCAAGTGACCA 3257
Db |||||
QY 2937 CGCTCTCATCTCCGCGCCACCACTGCTCATATTAAGTCTATCGGGCTCGTGGCGGCA 2996
Db |||||
QY 3258 CGCTCTGATCCGCGCCACCACTGCTGATCATCACTGTTCGGCGGTCTGTCGGCGCA 3317
Db |||||
QY 2997 TCTCGGACGCATCAACAAACGGGTACCACTCTCGGGGGCCCTCTTCGGCAAGCTCTTCT 3056
Db |||||
QY 3318 TCTCTACGCCATCAACAGCGGATACCACTCTCGGGGGCCCTCTTCGGCAAGCTCTTCT 3377
Db |||||
QY 3057 TCGCTTCTGGGTGATCTCCACCTCTACCGGTCTCTCAAGGGGTCTATGGGGCCGCA 3116
Db |||||
QY 3378 TCGCTTCTGGGTGATCTCCACCTCTACCGGTCTCTCAAGGGGTCTATGGGGCCGCA 3437
Db |||||
QY 3117 ACAGGACGCCACCGTGTGTGTCATCTGTCATCTCTGCGCTTCTCTCCCTGC 3176
Db |||||
QY 3438 ACGCACCCGACCATCTGTCGTCTGGGCGATCTCTGTCGGGTCTCTCTCTCTCTCT 3497
Db |||||
QY 3177 TCTGGGTGAGATGACCTTTCATCTGTCAGGACCAAGGGCCCGGACGTCAGGCGAGTGTG 3236
Db |||||
QY 3498 TGTGGGTTCGATGACCCCTTCAACCCCGGTCACTGGCGCGGATACCCAGAGCTGTG 3557
Db |||||
QY 3237 GCATCAATGCT 3248
Db |||||
QY 3558 GCATCAATGCT 3569
Db |||||
```

```
RESULT 3
AY103701 LOCUS AY103701 3788 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays P0120363 mRNA sequence.
ACCESSION AY103701
VERSION AY103701.1 GI:21206779
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3788)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S.,
Archur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3788)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1..3788
FEATURES
source
```

```
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:637005"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Query Match 39.1%; Score 1347.6; DB 3; Length 3788;
Best Local Similarity 65.7%; Pred. No. 3.9e-256;
Matches 2204; Conservative 0; Mismatches 1002; Indels 147; Gaps 11;

QY 30 CTGTGAGCACCTGAGGGTCCGAGGCGGAGAGCTAGCTAGCAGCCGCGGCTCCGCGC- 88
Db |||||
QY 137 CTGTGCCGCCGCCCGGGGTTCTGTGCGGAGCGAGATCCGGCGGGGCGGGCGGGGCC 196
Db |||||
QY 89 -GCCATGAGGCCAGCGCGGGCTGTGGCGGCTCGCACACCGGAACGAGCTGTGCT 147
Db |||||
QY 197 TGAGATGAGGCTAGCGGGGCTGTGGCGGCTCGCATTAACCGGAACGAGCTGTGCT 256
Db |||||
QY 148 GATCCGGGGCCA-CGAGGACCCCAAGCGCTTCGGGGG- - - - -CTGAGCGGGCA 195
Db |||||
QY 257 GATCCGGCGACCGCGAGTCGGGAGCGCGGGCGGGCGCGCGCGGGCGAGGC 316
Db |||||
QY 196 GGTGTGAGATATCGGCGGAGGTCGGGTACGGTGGAGCGGAGCTTCTGTCGC 255
Db |||||
QY 317 GCGTGTCCAGATATCGGCGGAGGTCGGGTTCGACGGGGAGCGCTTCGTGCG 376
Db |||||
QY 256 CTGCAACGAGTGGCGCTTCCCGGCTGTGGCGGCTGTACGAGTACGAGCGCGGAGGG 315
Db |||||
QY 377 GTGCAACGAGTGGCGCTTCCCGGCTGTGGCGGCTGTACGAGTACGAGCGCGGAGGG 436
Db |||||
QY 316 CACGAGAACTGCCCCCAAGTCAAGCGCTCAAGCGCTCAAGGGGAGCCCGAGGGT 375
Db |||||
QY 437 CTGCGGCGGTGCCCGAGTGCAGGACCGCGCTACAGCGCTCAAGGGTTCGCGCGGGT 496
Db |||||
QY 376 TCGCGGGAGCATGACGAGGAGGATCGAGCATCTGGAGCAGAGTTCAGATCGAGCA 435
Db |||||
QY 497 GCGTGGGGACGAGGAGGAGGACGCGCTCGAGCATCTGGAGGGCGAGTTCGCGCTGCAGGA 556
Db |||||
QY 436 CGAGAAATCAGCAGAGGAGGTCGGAGGCGCAACATCGAGAACAGCAGATCACCGAGGCGAT 495
Db |||||
QY 557 CGGCGCGGCCACGAGGACGACCGCGAGTACGTGCGCGAGTCCATGCTGCGGGGGCGAGT 616
Db |||||
QY 496 GCTGCAGCGGAGTGAAGTACGCGGAGGGGCGCGGAGCGAGCGGCGGCAACACACCCC 555
Db |||||
QY 617 GAGCTACGCGCGCGCGCGGCGAGCG- - -GCACCGGAGCTTCAACCCCGTCCCAAGGTGCC 673
Db |||||
QY 556 GCAGATCCGCGCCATATCACCAGCTCCCGCTCGGTGCGGTGAGCGGTGAGTTCGAT 615
Db |||||
QY 674 GCTCTTCAACAAACCGCCAGATGGTGTGATCATCCCGCGGAGCAGCAGCGGCTGTGCC 733
Db |||||
QY 616 TACCAACGGGTATGCGCCACGCGGAGGTCTGTCTTCTTCTGTCACAGCGCATCCATCCCTGA 675
Db |||||
QY 734 GTCCTATATGGCGCGCGCGCGGCGGCAAGAGGATCCACCGCTCCCTTCGCTGATTC 793
Db |||||
QY 676 CCCTGTGTCTGAGCCAGGGAGTGCACAGTGGGAGCAGAGAAAGAAAGA- - - - - 720
Db |||||
QY 794 CAACCTTCCAGTGCAACCGAGATCCATGAGCCCGCTCCAGAGGATCTGGCGCGCTACGGATA 853
Db |||||
QY 721 - - - - -AGTAGCTGGAAGGAGGAGGATGAGCAGCTGGAAGTCCCAAGCGGCGATCT- - - 771
Db |||||
QY 854 TGGCAGCGTGGCTTGGAGGAGAGATGGAGGGCTTGGAAAGCAGAGGAGGAGCGGCTGCA 913
Db |||||
QY 772 - - - - -CGGCGCGCGCGCGGATCCGAGAGATGACCGCGGAGTGGGCGGCGGAGTTCGAT 819
Db |||||
QY 914 GCATGTGAGGAGCGAGGGTGGCGGTGATTTGGGATGGGCGAGCATGCAGATCTGCCACTAAT 973
Db |||||
```


Db 3131 AG-----GAGTCTCGGAGCTGTACAGTTCAAGTGGACACCCCTGCTGTATACCCCCGA 3184
 Qy 2955 CCACGCTGTCATCATTAACGTCATCGCGGTGCTGGCCGGCATCTCCGACGCCATCAACA 3014
 Db 3185 CCACGCTCTCTGCTGAATTCATCGGGTGGTGGCCGGATCTCGAACGGATCAACA 3244
 Qy 3015 ACGGTACAGTCTCGGGGCCCCCTTCCTCGGCAAGCTCTTCTTCCGCTTCGTGGTCAATCG 3074
 Db 3245 ACGGTACAGTCTCGGGGCCCCCTTCCTCGGCAAGCTCTTCTTCCGCTTCGTGGTCAATCG 3304
 Qy 3075 TCCACCTCTACCCGCTCTCAAGGGGCTCATGGGGCCGAGAACAGGAGCCGCCCGTTG 3134
 Db 3305 TCCACCTGTACCCGTTCTCAAGGTCTGGTGGGAGGAGAACAGGAGCCGCCGATCG 3364
 Qy 3135 TTGTCATCTGTTCATCTCTGCTGCTCCATCTTCTCCCTGCTCTGGGTTCAGGATCGACC 3194
 Db 3365 TCATCGTCTGTTCATCTCTGCTGCTCCATCTTCTCCCTGCTCTGGGTTCAGGATCGACC 3424
 Qy 3195 CTTTCATCTGTACGACCAAGGGCCCGAGCTCAGGACGTGTGGCATCAATTGC 3247
 Db 3425 GTTCTCTGCCAAGAGACGCGCCGCTCTGGAGGAGTGTGGCTGGACTGC 3477

RESULT 4

LOCUS CL972423 3192 bp DNA linear GSS 21-SEP-2004
 DEFINITION OsIFCC022918 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.

ACCESSION CL972423.1 GI:52399375

KEYWORDS

SOURCE Oryza sativa (indica cultivar-group)

ORGANISM Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehartoideae; Oryzae; Oryza.

REFERENCE 1 (bases 1 to 3192)

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,

Wong, G. K. S., Deng, X. W. and Wang, J.

An analysis of transcriptional regulation of the rice genome and

its comparison to Arabidopsis

Unpublished (2004)

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

FEATURES Location/Qualifiers

1..3192

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Express Library"

/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 37.0%; Score 1272.6; DB 9; Length 3192;
 Best Local Similarity 64.7%; Pred. No. 2.7e-241;
 Matches 2069; Conservative 0; Mismatches 1029; Indels 99; Gaps 8;

Qy 147 TGATCCGGGGCCACGAGACCCCAAGCCGCTGCGGGCCGCTCAGCGGCGAGGTGTGGAGA 206

Db 2 TGGACGGGACGCGGATCCCGTGAAGTCTGGGAGGACACGGGAGGCGGCGTCCACA 61

Qy 207 TATCGGGGACGAGGTGGGCTCAGGTGGAGCGGACCTCTCGTCCCTGCAACGAGT 266

Db 62 TCTCGGGGACGCGGTGGGACGAGCGGGAGGCGGACGTGTTCGCCCTCGACGCTCT 121

Qy 267 GCGCTTCCCGGTGCGCGCCCTGCTACGAGTACGAGCGCGGAGGCGACGCAAACT 326
 Db 122 GCGCTTCCCGGTGCGCGCCCTGCTACGAGTACGAGCGCGCAAGGATGCGCAAGCTT 181
 Qy 327 GCGCCAGTGAAGACGGCTTCAAGCGCTTCAAGGGGAGGCGCGAGGTTGCCGGGAGC 386
 Db 182 GCGCCAGTGAAGACCGCAAGTACAAGCGCCCAAGGGGAGGCGCGGCGATCCGTGGGAGG 241
 Qy 387 ATGACGAG-----GAGGACATCGACGACCTGGAGCAGGATTCACATCG 431
 Db 242 AAGCGAGGATCTGATGCTGATGATGTCAGTGACTCAACTACCTGCATCTGCGAGTG 301
 Qy 432 ACGACGAGAAATCAGCAGGAGCAGCTGGAGGGCAACATGCAAGAACAGCCA-----GATC 484
 Db 302 CCGACCAAGACAGAAATTCCTGATAGGATGCGCAGTTGGCGCATGAATGCTGGGGGTG 361
 Qy 485 ACCGAGGCGATGCTGCAAGGAGGATGAGCTACGGGAGGCGCCGACGACGGCAGCGC 544
 Db 362 GTGAGATGTGCGCGCTCCCAAGTATGACAGTGGCGAGATCGGGCTCAAAATCTCGGG 421
 Qy 545 AACCAACCCCGCAGATCCCGCCCATCATCAACCGCT-----CCGCTCCGTCGCGG 596
 Db 422 AGAAATCCCTGGAGCTTCCCTGATCATATGATGTCTCTACCGGAAACATTTGCA 481
 Qy 597 TGACGGTGAAGTTTCCGATTAACCAACGGGTATGGCCACGGCGAGGTCTCTTCCCTTC 656
 Db 482 AGCGTGTCCATTCCTTATGTGAACCATTCACCAATCCATCAAGGGAGTTCTCTGGCA 541
 Qy 657 ACAAGCGATCCATCCGTACCTGTCTGAGCCAGGAGTGCAGTGGGACGAGAGA 716
 Db 542 GCATTGGAAATGTTTCCCTGGAAAGAAAGATGTGATGGCTGGAAACTGAAGCAGGACA 601
 Qy 717 AAGAAATGAGCTGGAAGGAGAGGATGACGACTGGAAGTCCAA---GACGAGCATCTCTCG 773
 Db 602 GAGCAATTCCTATGACCAATGGGACAGCATTTGCCCCCTTCTGAGGTGCGGAGTTGGTG 661
 Qy 774 GCGCGCGCGCATCCCGAAGACATGGAACGCGAGTGGCACTGAAACGACGAGCGAGGC 833
 Db 662 ATATCGATGCATCCACTGATTAATAATGGAAGATGCTTACTGAATGATGAAATCTGCC 721
 Qy 834 AGCGCTGTGAGGAAAGTGTGATCGGTGCGAGCAAGGTGAACCCGTACCGATGGTGA 893
 Db 722 AGCCTCTCTAGAAAGATTTCCCTTCTTCAATCAGAAATAAATCCCTACAGAAATGGTCA 781
 Qy 894 TCGTGTGCGTCTCGTTGTGCTCGCTTCTTCTCCGTCGCTACCGTATCTCGACCCCGTCC 953
 Db 782 TTGTTCTGCGTTGTTGTTCTAAGCATTTTCTTCACTACCGTATTTACGATCTCTGTCG 841
 Qy 954 CGGACGCCATCGGGCTGTGGCTCTCTCCATCATCTCGAGATCTGGTTCGCCATCTCT 1013
 Db 842 GCAATGCGTATCCGCTCTGGCTTTTATCTGTTATATGATGAGATTTGGTTTCTTGTCTCT 901
 Qy 1014 GGATCTCGACCAAGTTCCTCAAGTGTTCCTCCATGAGCGCGGAGAGCTACTCTGACCGCC 1073
 Db 902 GGATATGGATCAGTTCCTGAGGTGTTTCCAAATCAACCCGCGGAACTTACCTTATAGGC 961
 Qy 1074 TCTCCCTCAGTACGAGAGGAGGAGCGGTGCTGCTGCTGCGCGGTGAGACCTGTTTCG 1133
 Db 962 TGGCAATTTGAGGTATGACAGAGAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTTTG 1021
 Qy 1134 TGACACGGTGGACCCGCTCAAGGAGCGCGGTGGTGAGCCGCAACACCGTGTCTCTCA 1193
 Db 1022 TCAGTACAGTCGACCCCATGAAGGAACCTCTCTTGTACTGCGCAATACCGTGTGTCTCA 1081
 Qy 1194 TCCTCGCGGTAGACTACCCGTTGAGCAAGGTCTCTCTGCTAGCTCTCGACGACGCGCGT 1253
 Db 1082 TTCTGTCTGTGATTACCCAGTGGACAAAGGTCTCTTGTCTATGATATGATGATGTGCTG 1141
 Qy 1254 CGATGTGAGCTTCGAGTCTGCGGAGACGCGCGGAGTTTCGCGCGCAAGTGGGTGCCCT 1313
 Db 1142 CAATGCTGACTTTTGTATGCTATGGCTGAGACTTCAGAGTTTGTCTAGAAAGTGGGTTCAT 1201

1314	Qy	TCGTCAAGAAGTTCGGGATCGAGCCCGCGCCCGGGAGTTCTACTTCTCGCTCAAGGTGCG	1373
1202	Db	TCGTTAAGAAGTATAACATTTGAGCCAAAGAGCTCTCTGAGTGGTACTTCTCCGCAAAAATCG	1261
1374	Qy	ACTACCTCAAGGACAAAGTGTCAGCCCACTTCGTGTCAGGAGCGCGCGCCATGAAGAGAG	1433
1262	Db	ATTACTTGAAGACAAGTCCACCTTTCAATTTGTTAAAGACCGCCGCTGCCATGAAGAGAG	1321
1434	Qy	AGTATGAGGAGTTCAAGGTCGGAATCAACGCGCTGGTGGCCAAAGCCCATGAAGTGCCTGG	1493
1322	Db	AATATGAAGAAATTCARAGTTAGGATAAATGGCTTGTGTGCTNAGGCACAGNAAGTCCCTG	1381
1494	Qy	CAGAGGGTGGATCATGAAGAACGGCAACGCGTGGCCCGGGAAACAACAACCGCGGACCAAC	1553
1382	Db	AGGAGGGATGGATCATGCAAGATGGCAACCATGCCAGAGAAAACAATACTAGGAGCCATC	1441
1554	Qy	CCGCGATATCCAGGTGTTCTTGGGCCACAGCGCGCGCCACGACACCGAGGGGCAACGAGC	1613
1442	Db	CTGGAATGATTCAGGTTTCTTGGTCAACAGTGGTGGCTTGTATACCGAGGGTAAATGAGC	1501
1614	Qy	TGCCCCGCCCTCGTGTACGTCTCCCGTCGAGAAGCCCGCGGATTCACGACCAACAAGAAGG	1673
1502	Db	TTCCCCGCCCTGGTCTAGTATCTCGTGAAGAGCGTCTCGGGTTTCAGACCAACNAGAGG	1561
1674	Qy	CCGCGCCATGAACGCTCTGATTCGCGTCTCCGCGTGTCTGACCAACGCGCCATTCATGCG	1733
1562	Db	CTGGTGCATGAATGCTCTTGTTCGTGCTCAGCTGTGCTTACCAATGGACAGTACATGT	1621
1734	Qy	TCAACTTGGACTGTGATCACTAGATCAACAACAGCAAGGCCATCCGGGAGGCCATGTGCT	1793
1622	Db	TGAACTCTTGATTTGTGATCACTACAACAACAAGAGGCTCTCGGGGAAGCTATGTGCT	1681
1794	Qy	TCCTCATGAGCCCTCAGTGGCGCGGAAGTCTCTACGTTTCAGTTCCTCCGACAGAGTTGCG	1853
1682	Db	TCCTTATGGACCCAAACCTAGGAAGAGTGTCTGTTATGTTTCAGTTCCTCCCAAGAGTTGCG	1741
1854	Qy	ACGCGATCGACGTGCAACGATACGCTAAACAGGAACACCGTCTTCTTCGACATCAACA	1913
1742	Db	ATGGTATTGATPAGGAATGATCGATATGCGAACAGGAACACTGTGTTTTCGATATTAAC	1801
1914	Qy	TGAAGGGCTGGACGGGATCAAGGCCCGGTGTACGTTCGGGACAGGGTGGTGTTCGGCG	1973
1802	Db	TGAGGGGTCTTGATGGCATCCAGGACCAAGTTTATGTGGGAACCTGTTGTATTTCAACA	1861
1974	Qy	GCCAGGCGCTCTACGGCTACAACCTCCCAAGGGACCCCAAGAGG--CCCAAGATGGTGAC	2031
1862	Db	GAACCTGCTATATGGTTATGAAACCCCAATTAAGCAGAGAGAAGGGAAATTTCTTGT	1921
2032	Qy	CTGCGACTGTGCCGTGCTTTGGCGCGCAAGACGGAAAAACGCCCAAGGACGGCG---	2087
1922	Db	CATCACTTTGTGGGGCAGGAAGAAGCAAGCAAGTCAAGAAGAGATTCGGACAAAA	1981
2088	Qy	-----TGCCGG2093	
1982	Db	AGAACTCAAAACAAGCACTGGACAGTGTGTGCCAGTTTTTCAATCTTTGAAGATATAGAGG	2041
2094	Qy	AGGGCAACGCTGATATGGAGTAGTAGCGACAGGAGATGCTCATGTCACACATGAAC	2153
2042	Db	AGGGTGTGAAGGTGCTGGATTTCATGATGAGAAATCACTTCTTATGTCTCAAAATGAGCT	2101
2154	Qy	TCGAGAACGGTTCGGGCACTCCGCGGTTCGTCACTGCAACGCTGATGGAGGAAGCG	2213
2102	Db	TGGAGAGAGATTTGGCCAGTCTGACGGTTTGTGTGCTCCACTCTCATGGAATATGGTG	2161
2214	Qy	CGGTCCCTCTTCGTGAGACCCCGCGCTCTCTCAAGGAGGCCATTCATGTCTACAGCT	2273
2162	Db	GTGTTCTCTCAATCTGCAACCCGAAATCTCTTTTGAAGAAGCTATCCATGTGATAAGTT	2221
2274	Qy	CGGCTACGAGGACAGACCACTGCGGGCTGGAGCTGGGGTGTGATCTACGGGTGATCA	2333
2222	Db	GTGGCTATGAGGACAAAGACCGAAATGGGGAACTGAGATTTGGGTGTGATCTACGGTTCCG	2281
2334	Qy	CGGAGGACATCTCGACCGGGTTCAAGATGACACTGCGCGGGTGGCGCTCCGTGTACTGCA	2393

Db	2282	 CAGAAATATTCATCTGGATTCAAGATGCATGCGGAGGCTGGAGATCAATCTACTGCA	2341
Qy	2394	TGCCGAAGCGGGCGGGTTCAAAGGGGTCCGGCGCGATCAATCTATCGGACCGTCTCAAC	2453
Db	2342	TGCCCAAGCGCCAGCTTCAAGGGGTCTGCTCCTATCAATCTTTCAGATCGTCTTAAAC	2401
Qy	2454	AGTGCTCCGGTGGGCGCTGGGGTCCGTGAGATCTTTTCAGCCGGCAAGCCCGCTGC	2513
Db	2402	AAGTGCTCGGTGGGCACTGGTCTGTGTGAATCTCTTTTCAGTCGCGCAATGTGCCCAT	2461
Qy	2514	TGTACGGCTACAGAACGCGCAACTCAAGTGGCTGGAGGGCTTCGCTACATCAACACCA	2573
Db	2462	GGTACGGCTATGGA---GGACGCTTAAGTCTTGGAGAGATTTGCCATCATCAACACCA	2518
Qy	2574	CCATCTACCCCTTCACCTCGCTCCGCTGCTGCCTACTGCAACCCTCCCGCGCTGACC	2633
Db	2519	CTATTTATCCATTGACATCGATCCCGCTTCTCATACTGTGTGTTTGGCTGCTATCTGTT	2578
Qy	2634	TCCTCAGCGCAAGTTCATCATCGCTCGATAGACGTTGCGCAGGCTCTTCTTCAATCG	2693
Db	2579	TGCTCACTGGGAAGTTCATCATCCCAAGATTAGCAACTTTGCTAGTAAATTGGTTCACT	2638
Qy	2694	CCCTCTTCATGTCCATCTTCGCGACGGGCACTCTGCAGATCGCGTGAGCGGGGTGAGCA	2753
Db	2639	CTCTCTTCATTTCAATTTTGGCCACTGGTATCTTTGAGATGAGGTGAGTGGTGGCA	2698
Qy	2754	TCGAGGAGTGGTCGAGGAACGACGAGTTCTGGGTCACTCGCGGGGTGTCGCGCATCTCT	2813
Db	2699	TCGATGAGTGGTGGAGGAATGAACAGTTCTGGGGTTATTGGAGGTATATCTGCGCATCTTT	2758
Qy	2814	TCCGCGTGTGCAAGGCGCTCTCAAGGTCTCTCGCGGATCGACACAACTTCACCGTCA	2873
Db	2759	TTCCGCTCTTCCAGGGTCTCCTCAAGGTGCTGCTGGTATCGACACCAATTTCACTGTCA	2818
Qy	2874	CCTCCAAGGCCACCGCGCAGCAGGACGAGTTCGCGGAGCTCTACGCCCTTCAAGTGGGA	2933
Db	2819	CCTCAAGGCTTCTGATGAAGATGGTGAC---TTTGTGAGCTCTACATGTTCAAGTGGGA	2875
Qy	2934	CGACGCTCTCATCCCGCCACACAGGTGCTCATCATTAACGTTCATCGGCGTCTGTGGCG	2993
Db	2876	CAACGCTTCTCATCCCAACCGACCACTCTTGATCATTAACCTGGTCCGTGTTGTGCTGTG	2935
Qy	2994	GCATCTCCGACGCCATCAACAAAGGGTACCAGTCTTGGGGGGCCCTCTTCGGCAAGCTCT	3053
Db	2936	GTATCTCATACGGAATCAACAGCGGTACCAATCATAGGGACCGCTCTTTGGCAAGCTCT	2995
Qy	3054	TCATTGCGCTTCTGGGTCACTGTCACCTCTACCCGTTTCTCAAGGGGCTCATGGGGGCC	3113
Db	2996	TCATTGCGCTTCTGGGTGATTGTCCACTGTGTACCCCTTCTCTCAAGGGTCTTATGGGTCCG	3055
Qy	3114	AGAACGAGCGCCACCGTTGTCATCTGGTCCATCTGCTGGCTCCATCTTCTCTCC	3173
Db	3056	AAACCGCACTCCGACCATCGTTGTTGTTTGGGCAATCCCTCTTGGCTTCGATCTTCTCAT	3115
Qy	3174	TGCTCTGGGTTCAGATCGACCCCTTTATCTCGTCAGGACCAAGGGGCCGAGCTCAGGCAGT	3233
Db	3116	TGCTGTGGGTTCGATCGATCCATTACCAACCGGAGTCAACGCGCCAGATACCCAAACAT	3175
Qy	3234	GTGCGATCAATTTGCTGA	3250
Db	3176	GTGCGATCAACTGCTAA	3192

PRECIPITATION

AV104730	LOCUS	AV104730	3783 bp	mRNA	linear	HTC 16-OCT-2002
AV104730	DEFINITION	Zea mays PCO100501 mRNA sequence.				
AV104730	ACCESSION	AV104730				
AV104730	VERSION	AV104730.1				
KEYWORDS	HTC.	GI:21207808				
SOURCE	Zea mays					
ORGANISM	Zea mays					

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 3783)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

AUTHORS

Maize Mapping Project/DuPont Consensus Sequences for Design of

TITLE

Overgo Probes

JOURNAL

Unpublished (2002)

REFERENCE

2 (bases 1 to 3783)

AUTHORS

Coe, E.H.

JOURNAL

Direct Submission

COMMENT

Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSI, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers
1..3783
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 36.8%; Score 1266.4; DB 3; Length 3783;
Best Local Similarity 63.9%; Pred. No. 4.7e-240;
Matches 2096; Conservative 0; Mismatches 1056; Indels 126; Gaps 7;

QY 87 GCGGATGGAGGCGCAGCGGGGTGGTGGCGGCTGCACAAACCGGAAACGAGTGTGC 146
DB 181 GGGGGATGGAGGCGCAGCGGGGTGGTGGCGGCTGCACAAACCGGAAACGAGTGTGC 240
QY 147 TGATCCGGGGCCACGAGGA-----CCCAAGCGCGTGCAGCGGCTGCAGCGGGCAGGTGT 200
DB 241 TCATCCCGCGATGGCGAGCCAGGCGCGAAGCCCATGGACCGGAAACGCGCAGGTGT 300
QY 201 GCGGATATGCGGCGCAGAGTGGGTTCAGGTGGAACGCGGACCTTTCGTGCGCTGCA 260
DB 301 GCCAGATTGCGGCGCAGAGTGGGGGCGCAACCCCGACGCGGAGCCGTTCGTGCGCTGCA 360
QY 261 ACAGTCCGGCTTCCCGTGGCGGCGCTTACGAGTACGAGCGCGGAGCGCGGAGCGC 320
DB 361 ACAGTCCGGCTTCCCGTGGCGGCGCTTACGAGTACGAGCGCGGAGCGCGGAGCGC 420
QY 321 AGAACTCCCGCCAGTGAAGACGCGCTACAGCGCCCTCAAGGGGAGCGCGGAGGTTCGG 380
DB 421 AGAACTCCCGCCAGTGAAGACGCGCTTCAAGCGCTTCAAGGGGTGGCGGCGGTGCGG 480
QY 381 GGGACGATGACGAGGAGCATCGAACCTTGGAGCAACGAGTTCACATCGACGAGCA 440
DB 481 GGGACGAGGAGGAGCGCGGTGACGACCTTGGAGAACGAGTTCACCTGGAGCGCAAGC 540
QY 441 ATCAGCAGGCGAGTGGAGGCAACATGACGACGAGCGAGT----- 483
DB 541 ACAGTCCCGCTACCTCGCGAGTTCATGCTTCAACCGCCCATGAGTACGCGCGCGG 600
QY 484 ----CACCGAGGCGATCTGCAGCGCAGGATGAGCTACGCGAGGCGGCGCGCGCGG 539
DB 601 CCGACCTCGAGCGGCGTCCGCGAGCCATTCCACCCCATCCCAATGTTCCCTCTCACC 660

QY 540 ACGGCAACACACCCCGCAGATCCCGCCCATCATACCGGCTCCGCTCCGTGCGGTGA 599
DB 661 ACGGACAGATGGTGCATGACATCCCGCGGACACGACACGCTTGTGCTCCGTGCGGTG 720
QY 600 GCGGTGAGTTCCGATTAACCAAGGGGTATGGCCACGCGGAGGTCTCGTCTTCCCTGACAC 659
DB 721 GTGGCGGGGAGAGAGATTCAACCTCTCCGTACGCGGATCCCAACCTTCTGTCGAC 780
QY 660 AGCGCATCCATCCGTACCCCTGTGTGAGCCAGGAGTGCACAGTGGGACGAGAGAAAG 719
DB 781 CGAGGTCTATGGACCCCTTCCAAGGATCTCGCGCATATGGCTACGGGAGCGTAGCATGA 840
QY 720 AAGTGAGCTGGAAGGAGAGATGGACACATGGGAAGTCCAAGCAGGAGTCTCTCGCGCG 779
DB 841 AGGAGAGATGGGAGAGCTGGAAGCAGAAAGCAGGAGGAGTGCACACGAGGAGAACGATG 900
QY 780 GCGCGCATCCGGAAGACATGAGCCGCGTGGCAGTGAACGAGCAGCGGAGGAGCGCGC 839
DB 901 GCGGCGGAGATGATGGTATGATGAGATCTACCACTAATGGATGAAGCTAGACAGCAT 960
QY 840 TGTGAGGAAGGTGTGATCCGCTCGAGCAGGAGTGAACCCGTACCGGATGGTGTGCTGG 899
DB 961 TGTCCAGAAAGATCCCGCTTCTTCAAGCCAAATCAACCCCTATAGGATGATTATATAA 1020
QY 900 TGGGTCTGTGTGCTCGCCTTCTTCTCGGTACCGTATCTCTGCACCCCGTCCCGGAGC 959
DB 1021 TTGCGGTAGTGGTGTGTTGTGTTCTTCTTCCATACCGAGTGAATCATCCGCTGCTGAT 1080
QY 960 CCATCGGCGTGGCTCGTCTCATCATCTGCGAGATCTGTTCCGATCTCTTGGATCC 1019
DB 1081 CATTTGCTTATGGCTCATATCTGTGATCTGTGAAATTTGGTTGCCATGTCTTGGATTC 1140
QY 1020 TCACACAGTTTCCCAAGTGGTTCCCATCGACCGCAGAGCTACCTCGACCGCTCTCCC 1079
DB 1141 TTGACCAAGTTTCAAGTGGTTCCTATCGAGAGGGAACCTATCTTGAACGCGCTGAT 1200
QY 1080 TCAGGTACGAGAGGAGGAGCGGCTGCTGCTGTCGGCGGTGGACCTGTTCTGTGAGCA 1139
DB 1201 TAAAGTTTGAACAGGAGGCGATCTTCTCAACTCGCCCTGTCGATTTCTTTGTGAGTA 1260
QY 1140 CGGTGACCCCGCTCAAGGAGCGCGCTGTGTGACCGCAACACCGTCTCTTCCATCTCTCG 1199
DB 1261 CGGTGATCCCTTGAAGGAACCTTCCATTGCTACTGCTAATATCTGTTCTATCTATCTTT 1320
QY 1200 CCGTAGACTACCCCGTGGACAAGTCTCTGCTACGCTTCCGACGAGCGCGCTGCTGATGC 1259
DB 1321 CGGTGAGTTATCCAGTTGATAAGGTTTCAAGTGTCTGATGATGGTGTGCTGCCATGC 1380
QY 1260 TGAGTTTCGAGTCTGTTCGAGACGCGCGAGTTGCGCGCGAGTGGGTGCGCTTCTGCA 1319
DB 1381 TGACATTTGAAGCATTTGCTGAAACATCTGAAATTTGCAAGAAATGGGTTCTTTTCTGCA 1440
QY 1320 AGAAGTTCCGATCGAGCGCGCGCGGCTTCTGCTCAAGTCCGACTAC 1379
DB 1441 AAGATATAGCTTGAAGCTCTGCTCCAGTGTGTTCTTCAACAGAGATAGACTACC 1500
QY 1380 TCAGGACAAAGTTCAGCGCCACCTTGTGAGGAGCGCGCGCATCAAGAGAGAGTATG 1439
DB 1501 TGAAGACAAAGTGGCGCCAAACTTTGTTAGAAACGAGAGCAATGAAGAGAGATG 1560
QY 1440 AGAGTTCAAGGTCGGATCAACGCGTGTGGCAAGGCCATGAAGTGGCGGCAAGG 1499
DB 1561 AGGAAATCAAGGTGAGAAATCAATGCTTGGTTGTTAAAGCCCAAAAGTTCTCTGAGGAG 1620
QY 1500 GGTGGATCATGAAGGAGCGCACCGCGTGGCGCGGGAACAACACCCGCGACCCACCGCGCA 1559
DB 1621 GATGGAATGCAAGATGGAATCCATGGCCCGGAAATATGTTCCGTGATCATCTGGA 1680
QY 1560 TGATCCAGGTGTTCTTGGGCGCACAGCGCGCGCGCCACGACCCGAGGGCAACGAGTGGCCCC 1619
DB 1681 TGATTCAGGTTTCTTGGTCAAGTGGTGGCCATGATGTGGAAGGAATGAGGTGCTC 1740
QY 1620 GCCTGCTGACGTCTCCCGTGAGAGCGCGCGGATTTCCAGACCAACAGAGCGCGCGG 1679

1741 GATTGGTTTATGTTTCAAGAGAAACCGCGCTACAAACCATCAAGAGAGCGTGGTG 1800
1680 CATGAACCGCTGTGATTCGGCTCTCCGCGTGTGACCAACCGGCCATTATGCTCAACT 1739
1801 CTATGAATGCATTTGGTCCGAGTCTCTGCTGTACTAACTAATGCTCTTATTTGCTGAAC 1860
1740 TGGACTGTGATCACTACATCAACAAAGCAAGGCCATCCGGAGGCGCATGTGCTTCTCTCA 1799
1861 TGAATTTGTGATCACTATATCAATATAGTAAGGCTATTAAGGAAGCAATGTGTTTATGA 1920
1800 TGGACCTCTCAGGTTCGGCGGGAAGGTCTGCTACGTTCCGTCAGAGGTTTCGACGGCA 1859
1921 TGGATCCTTTGCTTGGAAATAAAGTTTGTCTATGTGCAAGTTTCTCAAGATTGATGGGA 1980
1860 TCGACGTGACGACGATAGCTTAACAGGAACACCGTCTCTTCGACATCAACATGAAG 1919
1981 TTAGTCGCCATGATGATATGCTTAACAGAAATGTTGCTTTTTCGATATCAACATGAAG 2040
1920 GCGTGCAGCGCATCCAAAGGCCGGTGTACGTGCGGACAGGTCGCTGTTCCGGCGCAGG 1979
2041 GTTTGGAATGATTCAGAGGCCCAATTTATGTGGGTACTGATGTGTTTTCAGAAAGCAGG 2100
1980 CGCTCTACGCTCAACACCTCCCAA----- 2004
2101 CATTATATGGCTACGATGCTCCCAAAGAAAGCACCATCAAGAACTTGCAACTGCT 2160
2005 ----- 2005
2161 GGCACAAAGTGGTGATTTGCTGTGCTGTTTGGTAAACAGGAAGACCAAGAAAGACCA 2220
2034 GCGACTGCTGCCGCTGCTCGCGCGCAAGAG-----CGAAACACGCCCAAGGACG 2084
2221 AGACCTCTAACCTAAATTTGAAGATAAAGAACTTTTAAAGAAAGGAATAAATCAAG 2280
2085 GCGTGCAGGAGGACCGCTGATP-----GGAGTAGATAGCGACAAG 2129
2281 CCCCTGCATATGCTCTTGGTGAATTTGAAGCTGCTCCAGGAGCTGMAAATGAAGAGG 2340
2130 AGATGCTATGCTCCCAATGATCTCGAAGCGGTTTCGGGACAGTCCGGCGGCTTCGTCA 2189
2341 CTAGTATGTAATCAACAGAAAGTTTGAAGAAAGAAATTTGGCCAGTCTTCAGTTTGTG 2400
2190 CGTGCAGCTGATGAGGAAGCGGCTCCTCTCTGTCGAGCCCGCGCTCTCTCA 2249
2401 CATCCACATCTTGAAATGGTGAAACCTGAAGAGTGCAGTCCAGCTTCTCTCTGA 2460
2250 AGGAGGCCATCCATGTCATCAGCTGCGGCTACGAGGAAGACCGACTGCGGCGCTGGAGC 2309
2461 AGGAAGCTATACATGTCATCAGCTGTGATATGAAGACAAACAGGCTGGGGAAGATA 2520
2310 TGGGTTGGATCTAGGGTGCATCAGGAGACATCTGACGGGTTCAAGATGCATGCTCC 2369
2521 TTGGTTGGATTTATGGATCAGTCACAGAATATTTTACTGGGTTTAAGATGCATGCTCC 2580
2370 GCGGTTGGCTCGCTGCTGCTGCTGATGCGAAGCGGCGGCTTCAAGGGGCTCGCGCGGA 2429
2581 ATGTTGGCGGTCAATTTACTGATACATCAACGGGCGGCTTCAAGGTTTCGACCTC 2640
2430 TCAATCTATCGGACCGCTTCAACAGAGTGTCTCGGTTGGGCGCTCGGTCCGTGAGATCT 2489
2641 TCAATCTTTCCGATCGTCTTCCACAGGTTCTTCGGTGGGCTCTTGGTTCAATTGAATTT 2700
2490 TCTTACCGCGGCACAGCCCTGCTGCTACGCTACAGAGCGCAACCTCAAGTGGGTGG 2549
2701 TCTTACAGAACCACTGCGCTCTCTGATGAGGTT---ATGGTGGTGGACTAAAGTTCTCG 2757
2550 AGGCTTTCGCTATACATCAACACCACTTACCCCTTTCACCTGCTCCGCTGCTCGCT 2609
2758 AAGGTTTTCGATCATTTAACTCCATCGTATACCTTGGACATATATCCGCTCTTGGCT 2817
2610 ACTGCACCTTCCCGCGCTGTGCTCTCTACCGGCAAGTTCAATCATCGCTCGATTAAGA 2669

2818 ATTGCACATTTGCTGCCATCTGCTGCTGACAGGGAATTTATCAGCCAGAGCTTAACA 2877
2670 CGTTCCGACGCTCTTTCTTTCATCGCCCTCTTTCATGCTCCATCTTCGGAAGCGGCTCTGG 2729
2878 ATGTTGCCAGCTGTGGTTCACTGCTTCTTTCATGCAATTTTGTACGAGATCTCTGG 2937
2730 AGATGCGGTGGAGCGGGGTGAGCATGAGAGGTGGTGGAGGAACGAGAGTTCCTGGGTCA 2789
2938 AATGAGATGGAGTGGTGGAGTGCATGATGATGAGTGGTGGAGAAACGAGCAGTTTGGGTCA 2997
2790 TCGGCGCGGTGTCGCGGCACTCTTCGCGCTGTCGAGGGCTGCTCAAGGTCCTCGCGG 2849
2998 TTTGAGGCGGTGCTTTCACATCTCTTTGCTGTGTTCCAGGAGCTTCTTCAAGGTCATAGCTG 3057
2850 GGATCGACACCAACTTTCACCGTCACCTCCAAGGCCACCGCGACGAGGAGCAGAGTTCG 2909
3058 GTGTAGACAGAGCTTCACTGTGACATCCAG-----GGCGGAGACGAGAGGTTCT 3111
2910 CCGAGCTTACGCTTCAAGTGGACCACTCTCTCATCCGCGCAACGAGCTGCTCATCA 2969
3112 CAGAGCTGTACACATTTCAAAATGAGCAGCCTTCTGATACCTCGGACAAACCTCTCTCTAC 3171
2970 TTAAAGTATCGGCGTGTGGCGGCACTCTCCGACGCCATCAACAGGGTACCACTCT 3029
3172 TGAATTTCAATGGAGTGTAGCTGGCATCTCCAATGCGATCAACAAACGGATATGAATCAT 3231
3030 GGGGGCCCTCTTTCGGCAAGCTCTTCTGCGCTTCTGGGTCATCTGCTCCACTTACCCGT 3089
3232 GGGGGCCCTGTTTCGGGAAGCTTCTTTCGATTTTGGGTGATCGTCCATCTTTACCCGT 3291
3090 TCTCAAGGGGTCTATGCGGCGCGCAGAACAGGACGCCACCGTTGTTGTCATCTGTTCCA 3149
3292 TCTCAAGGGTCTGTTGGAGGCGCAGAACAGGACGCCAACGATTTGTCATTTCTGGTCCA 3351
3150 TTCTGCTGCTCCATCTTCTCTCTCTGCTGCTGAGGATCGACCTTTCATCTGTCAGGA 3209
3352 TCTCTGCTGCTGATCTTCTCTGCTTGGGTCCGATCGACCGCTTCTTTCGGAAGG 3411
3210 CCAAGGCCCGGACGTCAGCAGTGTGGCATCAATTCG 3247
3412 ATGATGTCCTTCTGTTGGAGGAGTGTGGTCTGATTCG 3449

RESULT 6

AY112236
LOCUS CL1160_1 mRNA linear 3728 bp mRNA HTC 17-OCT-2002
DEFINITION Zea mays
ACCESSION AY112236
VERSION AY112236.1 GI:21216926

HTC.
KEYWORDS Zea mays
SOURCE Zea mays

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoidae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 3728)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,

Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of

Overgo Probes

JOURNAL Unpublished (2002)

REFERENCE 2 (bases 1 to 3728)

COE, E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat

Schnable, Iowa State, then clones may be requested from ZmDB:

www.zmdb.iastate.edu.

FEATURES
source

```
Location/Qualifiers
1. .3728
  /organism="Zea mays"
  /mol_type="mRNA"
  /db_xref="MaizeDB:630049"
  /db_xref="taxon:4577"
  /clone_lib="Maize Mapping Project/DuPont Cornsensus
  Library"
  /note="this sequence is part of a project of EST
  assemblies resulting from the application of public
  contigs to seed DuPont contigs; this resource was
  assembled by DuPont as part of a collaboration for the
  overgo addressing of BACs in conjunction with the Maize
  Mapping Project"
```

ORIGIN

Query Match	36.6%	Score 1259.4	DB 3	Length 3728	
Best Local Similarity	61.9%	Pred. No. 1.2e-238			
Matches 2020	Conservative	0	Mismatches 1098	Indels 147	Gaps 7
QY	133	GAACGACGTGTCGTGATCCGGGGCCACGAGGACCCCAAGCCGCTCGGGGCGCTGAGCGG	192		
Db	225	GAGCTCCCTCCATGAGGGCGACGGGAGCGGTGAAGNNNNNNNNNNNNNNNNNNNNNN	284		
QY	193	GCAGGTGTGGAGATATCGCGCAGCAGAGTCCGGGCTCACGGTGGACGGCACCTCTTCGT	252		
Db	285	NNNNNGTCCAGATCTCGCGGATGCGGTGGGCACCTACGCGCGAGGAGAGCTCTTCAC	344		
QY	253	CGCTCGAACAGAGTCGGGCTTCCCGGTGTGCGCGCCCTGTCTACAGATACGAGCGCGGGA	312		
Db	345	CGCTCGGAGGCTCTGCGGGTTCGCGGTGTGCGCGCCCTGTCTACAGATACGAGCGCAAGGA	404		
QY	313	GGGCACGCAAACTGCCCCACAGTCGAAGACGCGCTACAAGCGCTCAAGGGGAGCCCGAG	372		
Db	405	CGGCACACAGCTGNNNNNAGTGCNNNNNCAAGTACAGGCCACAAAGGGAGTCCAGC	464		
QY	373	GGTTGCCGGGGAAGATGACGAGGAGGACATCGACGAAGCTGGAGCAGAGTTCAACATC	430		
Db	465	GATCCGAGGGGAGGAGGACGATCTGATGCCGATGATGCTAGCGACTTCAACTACCC	524		
QY	431	-----GACGACGAGAACTCAGCAGAGGCGAGCTCGAGGGCAACATTCGAGACACCCAGAT	483		
Db	525	TGCATCTGGCAATGACGACCAAGACAGAGATTGCTGACAGGATGCGAGCTGGCGCAT	584		
QY	484	CACCGAGGCGAT-----GCTGCACGCGCAGGATGAGCTACGGGAGGGGCCCGACGACGG	537		
Db	585	GAATGCTNNNNNCAGCGGGGATGTTGGCGGCCCAAGTATGACAGTGTGAGATCGGGCT	644		
QY	538	CGACGGCAACACACCCGCGAGATCCCGCCCATCATCACGGGCTCCGCTCCGTGCGCGGT	597		
Db	645	TACCAAGTACGACAGTGGTGAGATCCCTCGGGGATACATCCNNNNNNNNNNNNNNNN	704		
QY	598	GAGCGGTGAGTTTCCGATTACCAACGGGTATGGCCACGGGAGGTCTCGTCTCCCTGCA	657		
Db	705	NATTTCCGGAGAAATCCTGTGTCTCCCTGACCATCATATGATGTCTCTACTGGGAA	764		
QY	658	CAACGCA-----TCCATCCGTACCCCTGTGTCTGAGCCAGGGAGTGCCAA	702		
Db	765	CATTGGCAGGCGCGCCCAATTCCTCTATATGAATCATTCATCAAAATCCGTCGAGGGAA	824		
QY	703	GTGGGACGAGAAAGAAAGATGAGCTCGAAGGAGGAGTGAACACATCGGAGTCCAAGCA	762		
Db	825	CTCTGTAGCGTTGGGAATGTTCCTCTGGAAGAGAGGGTTGATGGCTTGGAAATGGAAGCA	884		
QY	763	GGCA-----	767		
Db	885	GGACAGGGAAACAATCCCATGACGAATGGCAACGATTCCTCTCTGAGGGCCGGG	944		
QY	768	--TCTTCGGCGGCGCGCGATCCGAAGACATGGACGCGAGCTGGCAGTGAACGACGA	825		
Db	945	TGTTGGTGATATTGATGCACTCAACTGATTAACAACATGAAGATGCCCTTATTAAACGATGA	1004		
QY	826	GGCGAGGACCGCTGTTCGAGGAGGTGTGATCGCGCTCGAGCAAGGTGAACCCGCTACCG	885		

2085 TATTAACTTGAGAGGTCTTGATGGCATCCAAAGGACAGTTTATATGTGGCACCTGGCTGTGT 2144
Qy GTTCCGGCGCCAGCGCTCTACCGCTACCAACCTCCCAAGGGACCAAGAGGCCCAAGAT 2025
Db TTTTAAAGAGAGCTCTATATGTTTATGAGNNNNAAATTAAGCAAAGAGGGTGGTTT 2204
Qy GTTGACCTGCGACTGCTGCCCGTCTTCCGCCGCAAGAGAGCGAAACACGCCAAGGACGG 2085
Db CTTTGTATCATCTATGTGNN 2264
Qy GTTGCC----- 2091
Db GNNNTAG 2324
Qy GGAGGACCGCTGATATGGAGTAGATAGCAAGAGAGATGCTATGTCACCATGAA 2151
Db GGAGGAGATTGAAGGCGCTGGATTGATGATGAGAAATCACTTCTTATGTCCTCAATGAG 2384
Qy CTTTCGAGAAAGCGGTTCCGGCAGTCCGGCGGTTTCGTCACGTCGACGCTGATGAGGAAGG 2211
Db CTTTGGAGAGAGATTGGCCAACTCGAGCTTTTGTGCTCCACTCTGATGGAATNNN 2444
Qy CCGCGTCCCTCTTCTGTCGAGCCCGCCGCTCCTCAAGAGGCCATCATGTCATCAG 2271
Db NNTATC 2504
Qy CTTGGCTACGAGCAAGATTGAATGGGAACTGAGATTGGGTGGAATCATGTTCTGT 2564
Qy CAGGAGGACATCTCTACCGGGTTCAAGATGCACTGCGCGGGTGGCGCTCCGCTGACTG 2391
Db GACGGAAGATTTCTACTGGTTCAAGATGCAAGGAGGCTGGCGGTGATCTACTACTG 2624
Qy CATGCCAAAGCGCGCGCTTCAAGGGTCCGGCGCGATCAATCTATCGACCGTCTCAA 2451
Db CATGCCTAAGCGCGCGCTTCAAGGATCGGCTCCCATCAATCTCTCAGACCGCTCGAA 2684
Qy CCAAGTCTCCGGTGGCGCTGGGTCGGTCGAGATCTCTTCAGCGCGCACAGCCCT 2511
Db CCAAGTCTCCGGTGGCGCTCGGTTCAAGTGAATCTTTTCAGCGCGCATTTGNNNNT 2744
Qy GCTGTACGGCTACAAAGACCGCACTCAAGTGGCTGGAGCGCTTCGCCCTACATCAAC 2571
Db ATGTACGGGTAC--GGAGAGCGCTGAAGTTCTTGAGAGATTGCGCTACATCAAC 2801
Qy CACATCTACCGCTTCACTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2631
Db CACATCTACCGCTTCACTGCTCCGCTCCGCTCCTCAATTTACTGTATCTGCTGCTGCT 2861
Qy CTTCTCACCGGCAAGTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2691
Db CTTGCTCACGGGGAAGTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2921
Qy CGCCTCTTTCATGTCATCTTTCGCGAGCGGATCCTTGGAGATGGGTGGAGCGGGGTGAG 2751
Db CTCTCTTTCATCTCGATCTTCGCAAGGATATCTGAGATGAGGTGGAGCGGCTGGG 2981
Qy CATCGAGAGTGGTGGAGGACGAGCTTCTGGGTCAATCGCGCGCTGCTGCGGCTATCT 2811
Db CATCGAGAGTGGTGGAGGACGAGCTTCTGGGTCAATCGCGAGGATCTCCGCGCCACT 3041
Qy CTTTCCCGTGTGAGGCGCTGCTCAAGTCTCGCGGGATCGACACCACTTCAAGT 2871
Db CTTTCCCGTGTGAGGCGCTTCTCAAGGTCCTTCAAGGTCCTTCAAGGTCCTTCAAGG 3101
Qy CACCTCAAGGCGCAAGGAGGACGAGTTCGCGGAGCTCTACGCTTTCAGATG 2931
Db CACCTCAAGGCGCTCG--GATGAAGCGGAGCTTTCGCGGAGCTGTACATGTTCAAGT 3158
Qy GACACGCTCTCATCCCGGCCACACGCTGCTCATATTAAGTCAATCGGCTGCTGCTG 2991
Db GACGACACTTCTGATCCCGGCCACACCATCTGATCATCAACCTGCTGCGCGTGTGCTG 3218

Qy CGGCATCTCCGACGGCCATCAACACGGGTACAGTCTCTGGGGGCCCTCTTTCGGCAAGCT 3051
Db CGGCATCTCTAGGCCATCAACAGCGGGTACAGTCTGTGGGGTCCGCTCTTTCGGCAAGCT 3278
Qy CTTTTCGGCTTCTGGGTGATCGTCCACCTTACCGGTTCTTCAAGGGGCTCATGGGGG 3111
Db CTTTTCGGCTTCTGGGTGATCGTTCACCTGTACCGGTTCTTCAAGGGTCTCATGGGGT 3338
Qy CAGAACAGGACGCCACCGTTGTTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3171
Db GAGAACCGACCCCGACCATCGTGTGTTCTGGGCGATCTCTGCGCTGCTGCTGCTGCTGCT 3398
Qy CTTGCTCTGGGTGAGTCCAGTCCCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3231
Db CTTGCTGTTGGTTCGATCGATCCGTTTCCACCAACCGCTCACTGGCCGCGATCTCGAAC 3458
Qy GTGTGGCATCAATTCGTGAGCTGTT 3256
Db GTGTGGCATCAACTGCTAGGAGGT 3483

RESULT 7
AY103655
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY103655
Zea mays
AY103655
AY103655.1
HTC.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT

1 (bases 1 to 3897)
Arthur, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3897)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
1. 3897
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635496"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Query Match 34.9%; Score 1201.4; DB 3; Length 3897;
Best Local Similarity 62.7%; Pred. No. 3.5e-227;
Matches 2049; Conservative 0; Mismatches 1101; Indels 117; Gaps 7;

Qy 89 GCGATGGAGCCAGCGCGGCTGGTGGCGGCTCGCACACCGAGAGAGCTGGTGTG 148

Db 153 GCGATGGAGCGAGCGCGGGGTGGTGGCGGCTCCCAACGCGAACGAGCTCGTGTC 1212
QY 149 ATCCGGGGCCACGAGGACCCC-----AAGCGCGTGCGGGCGCTGAGCGGGCAGGTGTC 202
Db 213 ATCCGGCGGACGCGGATCCCGGGCGAAGCGCGCGGAGCAGAAACGGGACGGTGTGC 272
QY 203 GAGATATGGGCGGACGAGGTTCGGGTCAAGGTGGAGCGGACCTCTTGTCCTCCCAAC 262
Db 273 CAGATTGGCGGACGAGAGTTCGGCTTCGCCCGGGGACCCCTTGCTGGGTGCGAAC 332
QY 263 GAGTGGCGCTCCCGTGTGCGGCGCTGCTACGAGTACGAGCGCGGAGGCGACGAG 322
Db 333 GAGTGGCGCTCCCGTGTGCGGCGCTGCTACGAGTACGAGCGCGGAGGCGACGAG 392
QY 323 AACTGCCCGCCAGTCAAGACGCGCTACAGCGCTCAAGGGGAGCCCGAGGGTTCGCCGG 382
Db 393 AACTGCCCGCCAGTCAAGACTCGATACAGCGCTCAAGGGTTCGCAAGCTGTGACCGGT 452
QY 383 GACGATGACGAGGACATCGACGACTGGAGCAGAGTTCAACATCGACGACGAGAAAT 442
Db 453 GACGAGGAGGAGCGCGGTGATGACCTGGGACAAACGAGTTCAACTGGGACGCGCCATGAC 512
QY 443 CAGCAGAGCGAGCTGGAGGGCA-----ACATGCGAGACAGCCAGATCACCGAGCGATG 496
Db 513 TCGCAGTGTGGCGAGTCCATGCTCTACGGCCACATGAGCTACGGCGGTGGAGGTGAC 572
QY 497 CTGACCGGAGGATGAGCTACGGGAGGGGCCCGACGCGGAGCGGCAACAAACCCCG 556
Db 573 CTAATGGCGGCCCAAGCTTTCAGCTCAACCCCAATGTTCCACTCTCACCAACGGG 632
QY 557 CAGATCCCGCCCATCATACCGGCTCCCGCTCCGTCGGGTGAGCGGTGAGTTCCGATT 616
Db 633 CAAATGGTGGATGACATCCACCGGACGACGCGCTGTGCTCTTCTTCATGGGTGGT 692
QY 617 ACCAAGGATATGGCCACGGGAGGTCTCGTCTTCCCTGCAAGCGCATCATCCGTAC 676
Db 693 GGGGAAGAGAGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGG 752
QY 677 CTTGTGTCTGACCCAGGAGTCCAAAGTGGGACGAGAGAAAGAGTGAAGGAGGAG 736
Db 753 TCTATGGACCCATCAAGGATCTTGTCGATATGGGTATGGTATGTTGCTTGGAGGAA 812
QY 737 AGGATGGAGCATGGAAGTCCAAAGCAGGAGCATCTCGGC-----GGCGGCGCC 784
Db 813 CGGATGGAGAAATGGAAGCAGAGACAGAGAGGATGCACAGACGGGAAATGATGGTGT 872
QY 785 GATCCCGAAGACATGAGCCCGACGTGGCAGTGAACGAGCGGAGGCGAGCGGTGTCG 844
Db 873 GGTGATGATGGTGAAGTCTGATCTACCACTAATGGATGAAGAGAGCAACAACTGTCC 932
QY 845 AGGAAGGTGTGATCGCGTTCGAGCAAGGTGAACCGGTACCGATGGTGTATCGTGTGGT 904
Db 933 AGGAAATATCCACTTCCATCAAGCCAGATTAATCCATATAGGATGATTAATATTTCG 992
QY 905 CTCGTTGTGCTCGCTCTTCTTCGCGTACCGTATCTTGCAACCGCGTCCCGGACGCCATC 964
Db 993 CTTGTGGTTTTGGGGTCTTCTTCCACTACCGAGTGAATCGCGGTGAATGATGCAATT 1052
QY 965 GGGCTGTGCTGCTCCATCATCTGCGAGATCTGGTTCGCCATCTCTGGATCTCGAC 1024
Db 1053 GCTTTGTGGCGCATATCTGTTATCTGTGAAATCTGGTTGGCATGCTTGTGGATTCTTGAT 1112
QY 1025 CAGTTTCCCAAGTGTTCCTCATCGACCGGAGAGGTACCTCGACGGCTCTTCCCTCAGG 1084
Db 1113 CAATTTCCCAAGTGTTCCTTATGAGAGAGACTTACCTAGACGGGTGTCACTGAGG 1172
QY 1085 TACGAGAGGGAAGGGAGCCGCTCGTGTGTCGGGGTGGACCTGTGTGTGAGCAGGTTG 1144
Db 1173 TTCGACAGGAAGGGCAGCCATCTCAACTTGTCTCCAAATGATTTCTTTGTCAGTACGGTT 1232
QY 1145 GACCGCTCAAGGAGCGCGGTGGTGGACCGGCAACCGGTGCTCTCCATCTCTCGCCGTA 1204

Db 1233 GATCCCTTAAGGAACTCTTTGGTCAACAATACTGTTCTATCTATCTATCTCTTCGGTG 1292
QY 1205 GACTACCCCGTGGACAAAGTCTCTGTCTACGTCGTCGACGAGCGCGCTGATGTCGACG 1264
Db 1293 GATTATCTGTGTAAGGTTTCTTGCTATGTTTCTGATGATGGTGTGCTCAATGCTAACG 1352
QY 1265 TTGAGTCTGCTGTGGAGACGGCGAGTTCGGGCGAAGTGGTGGCTCTCTGCAAGAG 1324
Db 1353 TTTGAAGCATTAATCTGAAACATATGAAATTTGCAAGAAATGGGTTCCTTCTGCAACG 1412
QY 1325 TTGCGCATTCGAGCCCGCGCCCGGAGTTCTACTTCTCGCTCAAGGTTCGATCTCAAG 1384
Db 1413 TACATATTTGAACCTCGGCTCCAGAGTGGTACTTCCAAAGAGATAGATCTCTTGA 1472
QY 1385 GACAAAGTTCAGCCCACTTCTGTCGAGGCGCGCCCATGAAGAGAGATGATGAGAG 1444
Db 1473 GACAAAGTTCGAGCAAACTTTGTTAGGAGAGGAGCAATGAAGAGAGATGATGAGAA 1532
QY 1445 TTCAAGTTCGAGTCAACGCGCTGGTGGCGAAGGCCATGAAGTTCGCGCAGAGGGTGG 1504
Db 1533 TTCAAGTTCGAGTCAATGCTTGTGCTTGAAGCCAGAAAGTTCTTGAAGAGAGATGG 1592
QY 1505 ATCATGAAGGACGGCAGCGCTGGCGCGGAAACAAACCCCGACACCCCGGATGATC 1564
Db 1593 ACAATGCAAGATGGAACCCCTGGCTGGAACCAATGTTGCTGATCATCTCTGGAATGAT 1652
QY 1565 CAGGTGTTCTGGGCGCACAGCGGCGCCACGACACCGAGGCAACAGCTGCGCCGCTC 1624
Db 1653 CAGGTCTTCTTGGCAAGCGGAGGCTTGAATGAGGAAATGAATGCGCAAGATTG 1712
QY 1625 GTGTACGTCTCCGTGAGAGGCGCGGATTCAGCACCAACAGAGCGCGGCGCATG 1684
Db 1713 GTTTATGTTTCTAGAGAGAAACGACAGGCTATAACCATATAAGAAAGCTGTGCTATG 1772
QY 1685 AACGCTCTGATTGCGCTCTCCGCGTGTGACCAACGCGCAATCATGCTCAACTTGGAC 1744
Db 1773 AATGATTGGTCCGAGTCTCTGCTGTACTACCAATGCTCCATATTGTTAAACTTGGAT 1832
QY 1745 TGTGATCATCATACAAACAGCAAGGCCATTCGGGAGGCAATGCTCTCATGAGAC 1804
Db 1833 TGTGATCACTACATCAACAGCAAGGCTATAAGAGAGCAATGTGTTTTATGATGAC 1892
QY 1805 CTTAGGTTCGCGCGGAGGTCTGCTACGTTCACTTCCGCGAGAGTTTCGACGGCATCAG 1864
Db 1893 CTTTACTAGAAAGAGGTTTGTCTATGTAAGTTTCCCTCAAGATTTGATGGGATGAT 1952
QY 1865 GTGCACGACCGATACGCTAACAGGAAACACCGTCTTCTTCGACATCAACATGAAGGGCTG 1924
Db 1953 CGCATGACCGATATGCTAACCGGATGTTGCTTTTTGATATCAACATGAAAGTTG 2012
QY 1925 GACGGCATCCAAAGCCCGGTGTACGTGGGACAGGGTGTGTTCCGCGCGCAGCGCTC 1984
Db 2013 GATGGTATTCAAGGTCCAAATTAATGTTGTTACTGGATGTGATTTAGAAAGGCGAGCTTA 2072
QY 1985 TACGGCTACACCTCCCAAGGACCCAGAGGC-----CAAGATG 2026
Db 2073 TATGGTTATGATGCCCCCAAAAGAAAGCCACCATCAAGGACTTGAACCTGTGCGCC 2132
QY 2027 GTGACCTTCGAGCTGCTGCCGCTTGGCGCGAAGAGCGAAACACGCCAAGAGCGG 2086
Db 2133 AAGTGGTCTTTGCTGTTGCTGTTGGCAATAGGAAGCAANAGNAGACTACCAAAACC 2192
QY 2087 CTGCGCGAG----- 2095
Db 2193 AAAACAGAGAGAAAAAGTTATATTTTTTCAAGAAAGAGAGAACCAATCCCTGCATAT 2252
QY 2096 -----GGCACCGCTGATATGGAGTAGATAGGCAACAGAGATGCTCATG 2140
Db 2253 GCTCTTGGTGAATTTGACGAAGCTGCTCCAGAGCTGAGAAATGAAGGCGCGGTTATGTA 2312
QY 2141 TCCACATGAATCTCGAGAGCGGTTCGGGCGAGTCCGCGGCTTCTGTCAGCTCGACGCTG 2200
Db 2313 AATCAACAAAAATTAGAAAAAGAAATTTGGCCAAATCTTCTGTTTTGTATACACACTT 2372

Qy 2201 ATGAGGAGGCGCGTCCCTCTTCTGTCGAGCCCGCGCTCTCAAGAGGCCATC 2260
| | | | |
Db 2373 CTCGAGAATGTTGGAACCTTGAAGAGTCAAGTCTCTCTCTCTTTGAAGAAGCTATA 2432
| | | | |
Qy 2261 CATGTCATCAGCTCGGCTACGAGGACAGACCACTGGGGCTGGAGCTGGGTGATC 2320
| | | | |
Db 2433 CATGTCATAGTTTGGTTATGAAGACAAAGACAGACTGGGGAAGAGATTGGCTGATC 2492
| | | | |
Qy 2321 TACGGTTCGATCAGGAGGACATCTGACGGGTTCAAGATGCACTGCCGGGTGGCG 2380
| | | | |
Db 2493 TATGATCAGTTACAGAAGATTTCTAACTGTTTCAAGATGCAATTTGATGGGG 2552
| | | | |
Qy 2381 TCCGTGTAATCCGAGCGCGGCTTCAAGGGTTCGGCGCGATCAATCTATCG 2440
| | | | |
Db 2553 TCAATTTACTGCATACCTTAAACGGTTGCAATCAAGGTTCTGCACCTCTGAATCTTCA 2612
| | | | |
Qy 2441 GACGCTCAACACGAGTCTCGGTGGGCTGGGTCGCGAGATCTTTTACGCGG 2500
| | | | |
Db 2613 GATCGTCTTCAACAGGTCTTGGTGGGCTCTTGGTCTATTTGAGATCTTCTTCAGCAAT 2672
| | | | |
Qy 2501 CACAGCCCTCTGTACGGCTACAAGAAACGCAACCTCAAGTGGCTGGAGCGCTTGGC 2560
| | | | |
Db 2673 CATTTGCCCTCTTTGGTATGGGT---ATGGTGGCGGTCTGAATTTTGGAAAGATTTTCC 2729
| | | | |
Qy 2561 TACATCAACACCAACATCTACCCCTTCACTCGCTCCGCTGCTCGCTACTGCACCCCTC 2620
| | | | |
Db 2730 TACATCAACTCATCGTATCTCTTGGACATCTATTCCTCTTGGCTTACTGTACATTG 2789
| | | | |
Qy 2621 CCGCGCTGCTCTCTCACCGCAAGTTTCAATCATGCGTGCATAGACGCTTCCGACG 2680
| | | | |
Db 2790 CTTGCCATCTTTTATTGACAGGGAATTTATCTACTCCAGAGCTGAATATTTGGCCAGC 2849
| | | | |
Qy 2681 CTCTTCTTCATCGCCCTCTTCATGTCCATCTTCGCGAGCGGCATCTCGAGATCGGTGG 2740
| | | | |
Db 2850 CTGTGTTTCATGTACATTTTATCTGCATTTTGTCTACGACATCTCTAGAAATGATGG 2909
| | | | |
Qy 2741 AGCGGGGTGAGCATCGAGAGTGTGGAGGAACGAGCAGTTCTGGGTATCGCGGGGTG 2800
| | | | |
Db 2910 AGTGGTGTGGAATGTATGACTGTGTGGAGGAATGACGAGTCTGGGTCTATTGGAGGTG 2969
| | | | |
Qy 2801 TCCGGGATCTCTTCGCGCTGTGAGGGCTGTCTCAAGTCTCTCGCGGATCGACACC 2860
| | | | |
Db 2970 TCCTCACACCTCTTTGTGTGTTCAGGGAATCTCTCAAGGTCTATAGCTGTGATGATACA 3029
| | | | |
Qy 2861 AACTTTCACGCTCACTTCCAAAGCCACCGCGCAGGAGCAGAGTTTCGCGAGCTCTAC 2920
| | | | |
Db 3030 AGCTTTCACGCTGACATCAAG-----GTTGAGATGATGAGGAGTTCTCAGAGCTATAT 3083
| | | | |
Qy 2921 GCCTTCAAGTGGACACGCTCTCATCCCGCCACCGCTGTCTCATCATTAAGCTCATC 2980
| | | | |
Db 3084 ACATTCAAATGGACTACCTTATTGATACCTCTTACCACTTGTCTTATTGAATTCATT 3143
| | | | |
Qy 2981 GCGCTGTGGCGGATCTTCGAGCCCATCAACACGGGTACCACTCTCGGGGCCCCCTC 3040
| | | | |
Db 3144 GGTGTGGTGTGGCGGTTTCAATTCGATCAATAACGGATATGATCATATGGGGCCCCCTC 3203
| | | | |
Qy 3041 TTCGCAAGCTCTTCTTCGCTTCTGGGTCTATCGTCCACCTCTACCGGTTCTCAAGGGG 3100
| | | | |
Db 3204 TTTGGGAAGCTATCTTTGGATTTTGGGTGATGTGTCATCTTTATCCCTTTCTCAAGGT 3263
| | | | |
Qy 3101 CTCAATGGGCGCCAGAACAGACGCCCCACCGTTGTTGTCATCTGTGTCATTTCTGTGCC 3160
| | | | |
Db 3264 TTGGTTGGAAGGCAAAACAGACACCAACAGATTTGTCATGCTCTGTGTCATTTCTGTGCC 3323
| | | | |
Qy 3161 TCCATCTTCTCCGCTCTGGGTCAAGATCAACCTTTTCACTGTCAGGACCAAGGGCCG 3220
| | | | |
Db 3324 TCAATCTTCTCCGCTCTTGGGTTCGGAATGATCTCTTCTTTCGCAAGGATGATGGTCCG 3383
| | | | |
Qy 3221 GACGTCAAGGAGTGTGGCATCAATTGC 3247
| | | | |
Db 3384 CTTCTTGAGGAGTGTGGTTTGGATTGC 3410
| | | | |

RESULT 8
AY110415
LOCUS CL1166_1 mRNA 3898 bp linear HTC 17-OCT-2002
DEFINITION Zea mays
ACCESSION AY110415
VERSION AY110415.1 GI:21214824
KEYWORDS HTC
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3898)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3898)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1. 3898
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="maizemap:630057"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

Query Match 34.8%; Score 1197.2; DB 3; Length 3898;
Best Local Similarity 62.5%; Pred. No. 2.4e-226;
Matches 2026; Conservative 0; Mismatches 1130; Indels 85; Gaps 10;
Qy 89 GCGATGAGGCGCAGCGCGGCTGGTGGCGGCTCGCACCAACCGAAGAGCTGGTGTG 148
| | | | |
Db 186 GAGATGCGGCGCAACAGAGGAGTGGTGGCGGCTCGCACCAACCGAAGAGTGGTGTG 245
| | | | |
Qy 149 ATCCGGGCGCACGAGGACCC-----AAGCCGCTGCGGGCGCTGAGCGGGCAG 196
| | | | |
Db 246 ATCCGCGCACGAGGCGCGCGCGCGCGCTAAGCCACGAAGAGTGCAGATGGGCGAG 305
| | | | |
Qy 197 GTGTGCCAGATATCGCGGAGGAGTGGGCTCAGGTGGACGCGGACCTCTTGTGCGCC 256
| | | | |
Db 306 GTCTGCCAGATTTGCGGCGACTCTGTGGGCGTTTCAGCCACTGTGTGATGTCTTTGTGCC 365
| | | | |
Qy 257 TGCAACAGGTGGCGCTTCCCGCTGTGCGGCGCTGTCTACGAGTACGAGCGCGGAGGGC 316
| | | | |
Db 366 TGCATGAGTGTGCTTCCCTGTCTGCGGCTTGTCTATGATGAGGCGGCAAGAGGGG 425
| | | | |
Qy 317 ACGCAGAACTGCCCCAGTGCAGACGCGCTTCAAGCGCTTCAAGGGAGCCCGAGGTT 376
| | | | |
Db 426 AACCAATGCTGCTCTAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGATT 485
| | | | |
Qy 377 GCCGGGACGATGACGAGGAGGACATCGACGCTGGAGCAGGATCTCAACATGACGAC 436
| | | | |

Qy	2535	ACCTCAAGTGGCTGGAGCGCTTCGCTACATCAACACACCATCTACCCCTTCCACCTCGC	2594
Db	2701	GATTGAAGCTTTTGGAGGCGTGGCTTACATCAACACCAATTTTATCCAAATCACAATCG	2760
Qy	2595	TCCGCTGCTGCTACTGCAACCTCCCGCGCTGCTGCTCCTCACCAGGCAAGTTATCA	2654
Db	2761	TTCCGCTTATGCGCTATTTGTGTGCTTCTGCTATCTGCTTCTTTACCAATAAATTTATCA	2820
Qy	2655	TGCCGTGATPAGACGTTGCCAGCCTTCTTTCATCGCCCTCTTCATGTCATCTTCG	2714
Db	2821	TTCTTGAGATTAGCAATATGCTGGGATGTTCTTCATCTCTTTTCGCTCCANNNG	2880
Qy	2715	CGACGGGCATCCTGAGATGCGTGGAGCGGGTGAGCATCGAGAGTGTGGAGAACG	2774
Db	2881	CCACTGGTATATTGGAGCTTAGATGGAGTGGTGTGGCAATGAAGATTGGTGGAAATG	2940
Qy	2775	AGCAGTTCTGGGTATCGCGCGGTGTCGGGCATCTCTTTCGCGCTGTGAGGGGCTGC	2834
Db	2941	AGCAGTTTGGGTATTTGGTGGCACCTCTGCCCATCTCTTCCAGTGTTCAGGGTCTGC	3000
Qy	2835	TCAAGTCTCGCGGGATCGACCACTTACCGTCACCTCCCAAGGCCACCGCGCAGC	2894
Db	3001	TGAAGTGTGGCTGGGATGATACCACTTACAGTTTACCTCAAAGGCATC---TGATG	3057
Qy	2895	AGGACGACGAGTTTCCGCGAGCTTACGCTTCAAGTGGACCAACGCTCTCATCCCGCCA	2954
Db	3058	AGGATGGGACTTTGCTGAGCTATATGTTCAAGTGGACCAAGTTTGTCTCATCTCGA	3117
Qy	2955	CCAGCTGCTCATATTAACTCATGCGCGTGTGGCGGCAATCTCGAGGCCATCAACA	3014
Db	3118	CCACTGTTTGTCTATTAACTGCTGGGAATGGTGGCAGGAATTTCTGATGCCATTAA	3177
Qy	3015	ACGGGTACCACTGCTGGGGGCTCTTTCGCAAGCTCTTCTGCGCTTCTGGGTCTCG	3074
Db	3178	GTGGCTACCAATCTTGGGGTCCGCTCTTTGGAAGCTGTTCTTCTGATCTGGGTGNTCC	3237
Qy	3075	TCCACCTTACCCGCTTCTCAAGGGGCTCATGGGGCGCCAGAACAGACGCCACCGTTG	3134
Db	3238	TCCATCTGTACAGCTTCTCAAGGGTCTCATGGGAAGGACAAACGACACCAACATCG	3297
Qy	3135	TTGTCACTGCTCATCTGCTGGCTTCATCTTCTCCTGCTCTGCTGCTGAGTCAAGTCA	3194
Db	3298	TCAATGCTGCTTCCATCTTCTTGTGATCTATCTTCTCTCTGCTGCTGCTGAGTCA	3357
Qy	3195	CTTTCACTGCTCAGGAC---CAAGGCGCGGAGCTCAGGACGTTGGGATCAATTTGCTG	3251
Db	3358	CTTTCACTCCCGACACAGAAAGCTGCTGCTTGGGGCAATGTGGCTCAACTGCTGAT	3417
Qy	3252	C 3252	
Db	3418	C 3418	
RESULT 9			
LOCUS	AY110079		
DEFINITION	Zea mays CL1164_1 mRNA sequence.	mRNA	linear
ACCESSION	AY110079		HTC 17-OCT-2002
VERSION	AY110079.1		
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 3696)		
TITLE	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.		
JOURNAL	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
REFERENCE	Unpublished (2002)		
AUTHORS	2 (bases 1 to 3696)		
	Coe, E.H.		

TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.
FEATURES	Location/Qualifiers
source	1..3696
	/organism="Zea mays"
	/mol_type="mRNA"
	/db_xref="MaizeDB:630055"
	/db_xref="taxon:4577"
	/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
ORIGIN	
Query Match	32.9%; Score 1132.8; DB 3; Length 3696;
Best Local Similarity	60.2%; Pred. No. 1.4e-213;
Matches 1965; Conservative	0; Mismatches 1158; Indels 141; Gaps 9;
QY	155 GGCACACGAGGACCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGTGCGAGATATGCGGC 214
Db	260 GCGCGCAGCCACGAATTCGGGGAAGCATGTGCGGGCAGGTGTGCCAGATCTGCGGC 319
QY	215 GACGAGGTGCGGCTCAGGTGACCGGCACTTTCGTGCTGCGCTGCAACGAGTGGCGTTC 274
Db	320 GACGGCTGGGACACCGCGCGACCGGCACTTTCACCGCTTGCACGCTCTGCGGTTC 379
QY	275 CCGGTGTCGGCCCTGCTACGAGTACGAGCGCGGAGGCGCAGCAGAACTGCCCCCAG 334
Db	380 CCGGTGTCGGCCCATGCTACGAGTACGAGCGGCAAGGACGGCACCCAGGGCTGCCCGCAG 439
QY	335 TGCAAGACGCGCTCAAGCGCTCAAGGGGAGCCCGAGGGTTGCGGGGACGATGACGAG 394
Db	440 TGCAAGACTAAGTACAAAGCGCCAAAGGGAGCCACCACTACACGGTGAAGAAATGAG 499
QY	395 GAGGACATGACGACCTGGAGCAGCGATTCACATC-----GACGAGGAATTCAG 445
Db	500 GATGTGGATGCTGACGATGTGAGTGACTACCACTACCAAGCATCTGGCAACCCAGGATCAG 559
QY	446 CAGAGGCGAGCTGGAGGGCAACATGCAGA---ACAGCCAGATCACCGAGGCGATGCTGCA 501
Db	560 AAGCAAAAGATTGCTGAGAGAAATGCTCACTTGGCGGACAACTCACGTGGCAGTGATATT 619
QY	502 CGGCAGGAT-----GAGCTACGGGAGGGGCCCCGACGACGCGCGACGGCAACCAACA 551
Db	620 GGCCTGCTAAGTATGACAGCGGTGAAATTTGGGCATGGGAAGTATGACAGTGTGAGATC 679
QY	552 CCCCAGAGATCCCGCCCATCATCAGCGCT--CCCGCTCCGTGCGGTGAGCGGTGAGATT 609
Db	680 CCTCGTGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAGATTCCTGGAGCT 739
QY	610 TCCGATTACCAACGGGTATGGCCACGCGGAGGTCTCGTCTTCTTCCCTGCACAGCGCATCCA 669
Db	740 TCCCTGATCATATGATGTCTCTGTTGGGAACATTTGGCAGGCGTGGACATCAATTTCTCT 799
QY	670 TCCGT--ACCTGTGCTGAGCCAGGAGTGCCTCAAGTGGGACGAGAAAGAGTACG 727
Db	800 TATGTAATCATTTCTCCAAACCCATCGAGGGAGTTCTCCGTAGCTTGGCAATGTTTCA 859
QY	728 TGGAGGAGGAGGATGGACGACTGGAAATGCCA----- 759
Db	860 TGGAAAGAGAGGGTGGATGGATGGAAATGAAGGATAAAGGTGCAATTCCTATGACCAAT 919

QY 760 -----GCAGGCGATCCTCGGGCGGCGCGCGATCCCGAA 793
Db 920 GGAAACAAGCATTTGCTCCATCAGAAAGGCGTGAGTTGCTGATTAATGATGCTTCTACTGAT 979
QY 794 GACATGAGCGCGCGAGTGGCACTGAACGACGAGGCGAGCGCGCTGTGCGAGGAAGTG 853
Db 980 TATAACATGGAAGATGCTTACTGAATGATGAACTCGGCAACTCTATCTAGAAAAGTG 1039
QY 854 TCGATCGCGTGCAGCAAGGTGAACCCGTAACCGATGGTGAATGCTGCTCGTGTG 913
Db 1040 CCAATTCTTTCATCCAGAAATAATCCGTACAGATGGTCAATTGCTAGCTAGTTGCTGCT 1099
QY 914 CTCGCTTCTTCTCCGATACGATACGATCCTGACCCCGTCCCGAGCGCATCGGGTGG 973
Db 1100 CTATGCATATTCTTGGCTACCGTATCAACATCCTGTGAACAATGCAATATCCACTGTGG 1159
QY 974 CTCGCTCCATCATCTCGGAGATCTGTTGCGCATCTCTCTGATCCTCGACAGTTCCCG 1033
Db 1160 CTTTATCCGTCATATGAGATCTGTTGCTGTTGCTGCTGATTTTGGATCAGTTCCCA 1219
QY 1034 AAGTGGTTCCCATCGACCGGAGACGTACTCGACCGCTCTCCCTCAGGTAGCAGAGG 1093
Db 1220 AAGTGGTCCCAATCAACCGTGAACATACCTTGATAGACTGGCTTTAAGGTATGACCGA 1279
QY 1094 GAAGGGAGCGTCTGCTGCTCGCGGTGACCTGTTCTGTAGACACGCTGGAACCGCTC 1153
Db 1280 GAAAGGTGAACCATCTCAATTAGCTCTCTGTATANNNGTCAGTACTGTGATGCCAATG 1339
QY 1154 AAGGAGCGCGCTGTCACCGCAACCGTGTCTCCATCTCCGCGGTAGACTACCC 1213
Db 1340 AAGGAGCTCTCTTGTACATGCAATATCTGTGCTTTCCATCTTGTGCTGATATCCG 1399
QY 1214 GTGACAAGGTCTCTGCTAGCTCTCGACGACGCGCGGTGATGCTGACGCTTCGAGTCG 1273
Db 1400 GTTGACAAGGTATCTTGCTATGTTTCGATGATGAGCTGCTATGCTGACTTTTGATGCT 1459
QY 1274 CTGTGAGAGCGCGGAGTTCCGCGCAAGTGGTGGCTTCTGCAAGAAAGTTGCGGATC 1333
Db 1460 CTCTCTGAACCTTCAAGAGTTGCTAGAAAATGGGTTCGGTTCTGTGAAGAGTACAAATA 1519
QY 1334 GAGCCCGCGCGCGAGTCTTACTTCTCGCTCAAGTTCGACTACCTCAAGGACAAAGTG 1393
Db 1520 GAGCTTANGGCCCGGAAATGATCTTTGCTCAGAAAATGATTAATCTTGAAGAACAAGTT 1579
QY 1394 CAGCCCACTTCTGTCAGAGCGCGCGCCATGAAGAGAGATGATGAGAGTTCAAGTTC 1453
Db 1580 CAAACCTCATTTGTGAAGAAGCGCGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1639
QY 1454 CGGATCAACGCGTGTGGCCCAAGGCCATGAAGGTGCGCGGAGAGGTGATCATGAAG 1513
Db 1640 CGTATCAATGCTTGTAGCAAGCAACNNNNNGTTCCGAGGAGGATGATCATGCAA 1699
QY 1514 GACGCGACGCGTGGCGCGGAAACAACCGCGACCAACCGCGCATGATCCAGGTGTC 1573
Db 1700 GATGATACCTTGGCTTGGAAACAATATAGGAGCACTCTCGAAATGATTCAGGTTTC 1759
QY 1574 CTGGGCCACAGCGCGGCGCACACCGAGGCGCAACGAGCTGCGCGCTCGTGTAGTTC 1633
Db 1760 CTGGGTCAAGTGGAGGCTTGAAGTTGAAGCAATGAATCTCTCGTTGTTGTTATGTG 1819
QY 1634 TCCGCTGAGAAGCGCGCGGATTCAGACACAAGAGAGCGCGCGCATGAACGCTGTG 1693
Db 1820 TCTCGTGNNNNNCGTCTCGAATTCACATCAAGAAGGCTGGTGGCATGAATGCATT 1879
QY 1694 ATTCGCTCTCCGCGTCTGACCAACGCGCATTCATGCTCACTTGGATGTGATCAC 1753
Db 1880 GTTCGTGATCAGCTGCTTACTTAATGGGCAATACATGTTGAATCTTGAATTTGATGAC 1939
QY 1754 TACATCAACACAGCAAGGCGCATCCGGAGGCGCATGTCTTCTCATGAGCCCTCAGTCC 1813
Db 1940 TACATCAATATAGCAAGGCTCTTCGAGAAGCTATGTGCTTCTTATGGAACCCAAACCTA 1999

QY 1814 GGCGGAAGGTCTGCTAGTTCAAGTTCCCGCAGAGGTTTCGACGGCATCGACGTGCACGAC 1873
Db 2000 GGAAGGATGTCTGTATGTTCAATTTCTCAGAGGTTTGATGTGATTTGATAGGAATGAC 2059
QY 1874 CGATACGCTAACAGGAACACCGTCTTCTTCGACATCAACATGAAGGGCTGACGCGCATC 1933
Db 2060 CGATATGCAACAGGAACACCTGTGNNNNNCAATTAATCTTGAGAGGTTCTTGACCGCAT 2119
QY 1934 CAAGCCCGGTGTAGTCTGGGACAGGGTGCCTGTTCCGCGCCACAGGCGCTCTACGGCTAC 1993
Db 2120 CAAGGCGCAGTTATGTGGGAACCTGTTGTGTTTAACAGAACGCGCTTATATGTTAT 2179
QY 1994 AACCTTCCAAAGGACCCAGAGGCGCAAGATGTTGACCTGCGAGTGTGCTGCCGTGCTTC 2053
Db 2180 GAGCTCCAGTCAAGNNNNNNNNNGCCAGGCTTCTTCTCTGCTTTGTNNNNNAAGNNN 2239
QY 2054 GGCGCAAGAGCGGAACAACGCCAAG----- 2080
Db 2240 NNGAGCTCAAAATCTAAGAGAGCTCGGAAAAGAAAGTCAATAGACACGACAGACT 2299
QY 2081 -----GACGGGCTCGGAGGGCACCGCTGTATATGGGAGTAGAT 2119
Db 2300 TCTGTACCAGTATTTAATCTCGAAGATATAGAGGAAGGATTTGAAGTTCTCAGTTTGTAT 2359
QY 2120 AGCGACAAGAGATGCTCATGTCCCATGAACTTCGAGAAAGCGGTTCCGGGAGTCCGCG 2179
Db 2360 GATGAGAAATCGCTGATTTATGTTCTCAATGAGCTTTGGAGAGAGATTTGGCAGTCCAGT 2419
QY 2180 GCGTTCTGTCAGCTGAGCTGAGGAGGAGCGGCTGCTCTCTGTCGAGGCGCGCC 2239
Db 2420 GNNNNGTAGCTTCTCTGATGGAATATGTTGTTTCCAACTCTGCAACTCTGCAACTCCNNN 2479
QY 2240 GCGCTCTCAAGGAGGCGCATCCATGTCTCATGCTGCGGCTACGAGGACAAGACCGACTGG 2299
Db 2480 NNNNNNNNNNNNNNTATTCATGTCTCATCAGCTGTGCTATGAGGACAAAACTGACTGG 2539
QY 2300 GGGCTGAGCTGGGCTGATCTACGGGTGATCAAGGAGGACATCTGACGCGGTTCAAG 2359
Db 2540 GGAATCAGATTTGGGTGATCTATGTTCTGTTAAGAGACATTTCTCACCGGNNNNNN 2599
QY 2360 ATGCATGCGCGGCTGCGCTGCTGCTACTGCTGCGAGGCGGCGGCTTCAAGGG 2419
Db 2600 NNN 2659
QY 2420 TCGCGCGCATCAATCTATCGGACCGTCTCAACACAGGTGCTCGGTTGGGCGCTCGGGTCC 2479
Db 2660 TCTGCTCTCATCAACCTTTCGATCGTTTGAATCAAGTCTCGTGGGCTCTTGGTTC 2719
QY 2480 GTGAGATCTTCTTACGCGGACAGCGGCTGCTGTAAGGCTACAGAACGCGCACTC 2539
Db 2720 ATTGAAATCTTTTTCAGAGGCAATGTTCCCATATGTTATGCTGCT---ATGGAGGCGGCTT 2776
QY 2540 AAGTGGCTGAGGCTTTCGCTCATCAACACACCATCTACCCCTTCCCTCCTCGCTCCG 2599
Db 2777 AAATCTCGAGAGATTTGCTTATCAACAAATTTATCCACTCACATCAATCCG 2836
QY 2600 CTGCTCGCTACTGCACTCCCGCGCTCTGCTCTCTCCTCCTCAGCGCAAGTTTCAATGCG 2659
Db 2837 CTCTCTGCTACTGCAATTTGCCAGCAGTTTGTCTTCTCACTGGAAGTTTCAATCCCA 2896
QY 2660 TCGATAGACGTTTGGCAGCTCTTCTCATCGCCCTCTTCTCATGTCTCATCTTCCGAG 2719
Db 2897 AAGATTAGTAACCTAGAGAGTGTGTTGTTTATATCGCTCTTTATCTCAATCTTGGCACT 2956
QY 2720 GGCATCTGAGAGTCCGTTGAGGCGGTGAGCATCGAGGATGAGGAGGAGGAGGAGGAG 2779
Db 2957 GGTATCTCTGNN 3016
QY 2780 TTCTGGGTCTATCGCGGCTGTCCGCGCATCTCTTTCGCGCTGCTGACGAGGCTGCTCAAG 2839
Db 3017 TTCTGGGTCTATGTTGTTATTTTCGCGCATTTATTTTCGCGCTCTTCCAGGCTCTCTGAG 3076
QY 2840 GTCTCTCGCGGATTCGACACCAACTTCACTGCTCACTCCAGGCCACCGGCGACGAGGAC 2899

```
Db 3077 GTGCTTGCTGGTATCGACGAGCTTCACTGTCACTCTTAAGGCCAC---TGACGAAGAA 3133
Qy 2900 GACGAGTTGCCGAGCTTACGGCTTCAAGTGAGCACCGTCTCTATCCGGCCACACG 2959
Db 3134 GGTGATTTTGGCGAGCTTACATGTTCAAGTGACAAACGCTTCTGATCCCAACCACT 3193
Qy 2960 CTGCTCATCATTAACGTCATCGGCGTCGCGCGGCGGATCTCCACGCCATCAACAACGG 3019
Db 3194 ATTTGATCATCAACCTGTGCGCGTGGTGGCATTTCTACGCAATCAATACGGT 3253
Qy 3020 TACCAGTCTTGGGGGCCCTTCTCGCAAGCTCTTCTTGGCCCTTCTGGGTCTATCGTCCAC 3079
Db 3254 TACCAGTCTGCGGAGCTCTTTTCGGGAAGCTCTTCTTGGCTTCTGGGTGATGTCAC 3313
Qy 3080 CTCTACCGCTTCTCAAGGGGCTCATGGGGCGCGAGAACAGAGCGCCACCGTGTGTGTC 3139
Db 3314 CTGTACCGCTTCTCAAGGGGCTCATGGGGGAGAGAACAGAGCGCCACCGTGTGTGTC 3373
Qy 3140 ATCTGGTCCATTCTGTGGCTCCATCTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3199
Db 3374 GTCTGGGCTATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3433
Qy 3200 ATCGTTCAGGACCAAGGGCCCGGAGCTCAGGACGAGTGCGCATCAATTTGCTGAGCTGTTAT 3259
Db 3434 ACCACCGGCTACTGCTCTGATATCGCGAATGTGGCATCAACTGCTAGGATGAGCTG 3493
Qy 3260 TAAGGTTCAAAATCTGAGAGCTTG 3283
Db 3494 AAGATAGTTAAAGAGTGGAACTAG 3517
```

RESULT 10

```
CNS0A1ML
LOCUS
DEFINITION
  Arabidopsis thaliana Full-length cDNA Complete sequence from clone
  GSLTGH562A06 of Hormone Treated Callus of strain col-0 of
  Arabidopsis thaliana (thale cress).
ACCESSION
  BX832166
VERSION
  BX832166.1 GI:42459101
KEYWORDS
  HTG; GSLT cDNA.
SOURCE
  Arabidopsis thaliana (thale cress)
  Arabidopsis thaliana
  Arabidopsis thaliana
  Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
  1 (bases 1 to 3911)
  Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
  Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
  Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
  Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
  A Combined Approach to Evaluate and Improve Arabidopsis Genome
  Annotation
  Unpublished
  2 (bases 1 to 3911)
  Genoscope.
JOURNAL
  Direct Submission
REFERENCE
  Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  The sequences are based on single pass reads.
  Life Technologies (a division of Invitrogen) members carried out
  full-length libraries construction : Temple G.
  Genoscope members carried out sequencing and annotation : Castelli
  V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
  Schachter V., Weissenbach J., Salanoubat M.
  URV INRA : Clepet C., Caboche M.
  Annotation is based on the June 2003 version of the Arabidopsis
  genome released by MIPS (Munich information center for Protein
  Sequences). 5 prime and 3 prime are assembled with Phrap.
  http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full
  length
  http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
```

QY 896 GTGGTGGCTCGTTGGTCTCGCTTCTTCTCCGTACCGTATCTCTGACACCCGTCGG 955
DB 1265 GTGCTACGTCCTTGTGATCTTGGTCTCTTCTTCTACCTACCGTATCTTACACCCGTCAAA 1324
QY 956 GACGCCATCGGCTGTGGCTCGTCTCCATCATCTCGAGATCTGGTTGCCATCTCCCTGG 1015
DB 1325 GATGCATATGCTTGTGGCTTATCTGTATATATGAGATATGTTGCTGTTTCAATGG 1384
QY 1016 ATCTCTGACCAAGTTCCTCCAAAGTGTTCCTCATCGACCGGACGCTACCTCGACCGCTC 1075
DB 1385 GTTCTTGTATCAGTTCCTTAAATGGTACCTATCGAGCGGAGAAACGCTACTTGGACCGACTC 1444
QY 1076 TCCCTCAGTACGAGGGAAGGGAGCGCTGCTGCTCTCGCGGTGGACCTGTTCGTG 1135
DB 1445 TCRTTAAGATATGAGAAAGGGAACCGTCGGGACTATCCCTCTGGATGATATTTGTT 1504
QY 1136 AGCAGGTGGACCCGCTCAAGGAGCGCGCTGTGTGACCGCCAAACACCGTGTCTCCCATC 1195
DB 1505 AGTACAGTGGATCCATTGAAGAGCCTCGCTTATTACTGCAATACTGCTTGTCTATT 1564
QY 1196 CTCGGCGTAGACTACCCCGTGAACAAGTCTCTCTAGTCTCCGACGACGCGCGTGG 1255
DB 1565 CTTGCTGTGATATCTCTCGATAAAGTGTCTGTTACGTATCTGATGATGGTCTGCT 1624
QY 1256 ATGCTGACGTTTCGAGTCTGTCGGAGACGCGCGAGTTCGCGGCAAGTGGGTGCCCTTC 1315
DB 1625 ATGCTTACTTTCCGAAGCTCTTTCTGAGACCGCTGAATTCGCAAGGAAATGGGTTCCTTTC 1684
QY 1316 TCGAAGAAGTTCGGCATCGAGCCCGCGCCCGGAGTTCTACTCTCGCTCAAGTCTCGAC 1375
DB 1685 TCGAAGAATAATGTTGATGAGCTGTGCTCCGAAATGGTATTTCTGCCATAAATGGAC 1744
QY 1376 TACCTCAAGGACAAGTTCAGCCCACTTTCTGTCAGGAGCGCGCGCCATGAAGAGAGAG 1435
DB 1745 TACTTGAAGATAAAGTTTATCCCGCATTTGTTAGGAGCGCGAGCCATGAAGAGAGAT 1804
QY 1436 TATGAGGATTCAGAGTCCGATTCACGCGCTGTGTGGCCAGGCCATGAAGTGCAGCA 1495
DB 1805 TATGAAGAAATTCAAAGTAAAGTCAATGCTTTAGTAGCAACAGCAAGAAAAGTGTCTGAA 1864
QY 1496 GAGGGTGGATCATGAAGAGCGGACGCGTGGCCCGGGAACAACACCCGCGACCAACCC 1555
DB 1865 GATGGGTGACATTCGACAGCGGTACACTTGTGGCCCGTATATAGTGTGGAGATCATCT 1924
QY 1556 GGCATGATCCAGGT- GTTCTTGGGCCACAGCGCGGCCACGACACCGAGGGCAACGAGCT 1614
DB 1925 GGCATGAATCAGGTCTTTCTTGGAAAGTGACCGTGTCTGTGATGTCGAAACAAACGAGTT 1984
QY 1615 GCCCGCCTCGGTGATAGTCTCCGCTGAGAGCGCCCGGATTCAGACCAACAAGAGGC 1674
DB 1985 GCCTCGATTAGTTACGTTTCTCGTGAGAGAGACCCGGATTTGATCACCATAAGAGGC 2044
QY 1675 CGCGCCATGAACGCTCTGATTCGCTCTCGCCGCTGTGACCAACGCGCCATTCATGCT 1734
DB 2045 TGGAGCTATGAAATCCCTGATACGAGTCTCTGGGGTCTTATCAAAATGCTCTCTTCT 2104
QY 1735 CAACTTGGACTGTGATCACTACATCAACAACAGCAAGGCATTCGCGGAGGCCATGTGCTT 1794
DB 2105 GAATGTCGATTTGTGATCACTACATCAACAATAGCAAGCTCTTAGAAGCAATGTGTTT 2164
QY 1795 CCTCATGACCTCATGTTGCGCGGAGAGGTCTGTACGTTTCAAGTTCGCGAGAGGTTTGA 1854
DB 2165 CATGATGATCCTCATGTCAGGAAGAAATCTGTTATGTTTCAAGTTTCCCTCAAGGTTTGA 2224
QY 1855 CGGCATCGACGTGACGACCGATAGCTAACAAGGAACCGTCTCTTCTTCGATCAACAT 1914
DB 2225 TGGGATGATAGGACGATCGATATCTCAATCGCAATGTGTGTTCTTTGATATCAATAT 2284
QY 1915 GAAAGGGTGGACGGCATCCAAAGGCCGCGTGTACGTCGGGACAGGGTGGTGTTCGGCG 1974
DB 2285 GAAAGGTTTGGATGGGCTACAGGGCCCTATACGTCGGTACAGGTTGTGTTTTCAGGAG 2344
QY 1975 CCAGGCGCTCTACGGCTACAACCCCTCCCAAGGGACCCCAAGAGGCCCAAGATGGTGACCTG 2034

DB 2345 GCAAGCGCTTTACGGAATTTGATGCAACGAAGAAGAGGCCCACTAAGACATGCAA 2404
QY 2035 CGACTGCTGCCGCTGCTTCGGCC----- 2057
DB 2405 TTGCTGGCCAAATAGGTGTCTCTATATGTTTGGTTCAAGAAGAATCTTAAGCAAGAC 2464
QY 2058 -----GCAAGAAGCGGAACAACGCAAGCA 2082
DB 2465 AGTGGCTCGGATAAAGAAGAAGAAATAGGAACGCTCAAGACGATCCACGCAATTAGA 2524
QY 2083 CGGGCTCGCGGAGGACCGCTGATATGGGAGTAGTAGCACAAGGAGATGCTCATGTC 2142
DB 2525 AAATATCGAAGAGGCGCGCTCACTAAAGGTTCTAAACGTGAACAGTCAACCGAGGCAAT 2584
QY 2143 CCATATGAATCTCGAGAACGCTTCGGGACGCTCCGCGCGTTCGTCAGCTCGACGCTGAT 2202
DB 2585 GCAATGAAGTTGGAGAAGAAATTTGGGCACTCTCTCTGTAATTTGTTGCACTCGCGGAT 2644
QY 2203 GGAGAAAGCGCGGTCTCTCTCGTCAAGCCCGCGCGCTCTCAAGAGAGGCCATCA 2262
DB 2645 GGAGAAATGGTGGATGGCTAGAAAACGCAAGCCCGCTTGTCTGCTTAAAGAAGCCATCCA 2704
QY 2263 TGTATCAGCTGCGCTACGAGGACACGACCTGGGGGCTGGAGCTGGGGTGGATCTA 2322
DB 2705 AGTCATTAGTTGCGGATATGAAGATAAACTGAAATGGGGAAGAGAGATTGGGTGGATCTA 2764
QY 2323 CGGGTGCATCACCGAGGACATCTCGACGGGTTCAAGATGCATTCGCCCGCGGTGGCGTCT 2382
DB 2765 TGGTCTCTTACCGAGATATCTTACGGGTTTTAAGATGCATCTCATGGTTGGAGATC 2824
QY 2383 CGTGTACTGCAATGCCAAGCGGCGGTTCAGAGGTCGCGCGGATCAATCTATCGGA 2442
DB 2825 TGTATTTGTACACAAAGTTAGCGGCTTTCAAAGGATCAGCTCCAATCAATCTTTCGA 2884
QY 2443 CGCTCTCAACAGGTGCTCCGGTGGGCTGGGTCGGTCAGATCTTCTTCAGCCGCA 2502
DB 2885 TCGTCTCCATCAAGTCTTCGATGGGCGCTTGGGTGCGTTGAGATTTTCTTGAAGTGGCA 2944
QY 2503 CAGCCCCCTGTGTACGGCTACAAGAACGGCAACCTCAAGTGGCTGGAGCGCTTCGCCCTA 2562
DB 2945 TTGTCTATTTGGTATGGTT----ATGGAGGTGGGTTGNAATGGCTTGAAGCGGTGTCTCTA 3001
QY 2563 CATCAACACCACTACTACCCCTTCACTCGCTCCGCTGCTGCTACTGCTACTGACCCCTCC 2622
DB 3002 CATTAACCTCTGTGTTTACCGTGCACCTCTCTACCGCTCATCGTTACTGTCTCTCTCC 3061
QY 2623 CGCGCTGCTGCTCTCAACCGCAAGTTCACTATCGCGCTCGATTTAGCAGCTTTCGCGAGCT 2682
DB 3062 TGCCATCTGTCTCTCACTGGAAAATTCATCGTTTCCGAGATTAGCAACTATGCGAGTAT 3121
QY 2683 CTTCTTCATCGCCCTCTTTCATGTCCATCTTCGCGACGGGCACTCTGGAGATGCGGTGGAG 2742
DB 3122 CCTCTTATGGCGCTCTTCTCGTGAATGCAATTAACGGGTATTTCTGAGATGCAATGGGG 3181
QY 2743 CGGGGTGAGCATCGAGGAGTGGTGGAGAAACAGCAGTTCT- GGGTCACTCGCGCGCTGT 2801
DB 3182 CAAAGTTGGGATCGATGATTTGGTGGAGAAACGAAACAGTTTTTGGGGTCACTTGGAGGTGTT 3241
QY 2802 CCGGCACTCTCTTCGCGCTGTGAGGGCTGTCTCAAGGTCCTCGCGGGATTCGACCA 2861
DB 3242 CTGGGCACTCTGTTTCTCTTCTCAAGGTCCTCAAGGTTCTTCTGCTGCTGTCGACACTA 3301
QY 2862 ACTTCACCGTCACTCAAGGCCACCGCGACGAGGACGAGTTCGCGCGAGCTCTACG 2921
DB 3302 ACTTCACAGTCAATCAAAAGC-----AGCTGATGATGGAGAGTTCTCTGACCTTTTACC 3355
QY 2922 CTTTCAAGTGGACACGCTCTCTCATCCGCCACCAACGCTCTCTCATTAACGTCATCG 2981
DB 3356 TCTTCAATGACCTTCACTTCTCATCCCTCCCAATGACTCTACTCATCATAAACGTCATG 3415
QY 2982 CGGTGTTGGCGCGCATCTTCGACGCCATCAACAAACGGGTACCAAGTCTCTGGGGGCCCTCT 3041

Db 1206 TGGAGCTTAACATTATCATTAAGTTGCTGTGGCGGAAGAAAAAGAAAGGACAAGAGCT 1265
QY 2050 CTTTCGGCCCAAGAACCGGAACAACGCCAAGACCGGGCTGCCGAGGGCACCGCTGATAT 2109
Db 1266 ATATTGATTCCAAAACCGTGATATGAGAGAACAGAAATCTTCCTCGGCTCCCATCTTCAACA 1325
QY 2110 GGGAGTAGTAG-----CGAAGAGAGATGCTCATGTGCC 2144
Db 1326 TGAAGATATAGAAGAGGATTTGAAGGTTACGAGGATGAAGAGTCACTGCTTATGTCTC 1385
QY 2145 ACATGAACCTTCAGAAAGCGTTTCGGCAGTCCGCGCGTTCTGCTCAGCTCGACGCTGATGG 2204
Db 1386 AGAAGAGCTTGAGAAACGTTTGGCCAGTCTCCAAATTTTATGCAATCCACTTTATGA 1445
QY 2205 AGAAGGCGGCTCCTCTCTCGAGCCCGCGCTCTCAAGGAGGCAATCCATG 2264
Db 1446 CTCAAAGTGGCATACCCCTTCAACAAACCCAGGTTCCCTGCTAAAGGAAGCTATACATG 1505
QY 2265 TCATCAGCTGCGGTACGAGGACAAGACCGACTGGGGCTGGAGCTGGGGTGGATCTAG 2324
Db 1506 TCATTAGTTGTGGATATGAGGATAAACAAGAAATGGGGAAGAGATCGGATGGATATATG 1565
QY 2325 GGTGCATCAGGAGACATCTGACGGGTTCAAGTGCACCTGCCGCGGGTGGCGCTCG 2384
Db 1566 GCTCTGTTACTGAAGATATTTAACTGTTTCAAGATGATGCAAGAGGTTGGATCA 1625
QY 2385 TGTACTGCATCGAAGCGGGCGGTTCAAGGGGTGCGCGCGATCAATCTATCGGACC 2444
Db 1626 TCTACTGCATGCCACTTCGGCTTCTTCAAGGGTTCGCTCCGATTAATCTTCTGATC 1685
QY 2445 GTCTCAACAGGTGCTCGGTGGGCGTGGGTGCTGAGATCTTCTTACAGCGGCA 2504
Db 1686 GTCTCAACCAAGTTGATCGCTGGGCTCTGTTGCTGAAATTTACTTAGACACACT 1745
QY 2505 GCGCCCTGCTGACGCTACAGAAACGCAACTCAAGTGGCTGAGCGCTTCGCTACA 2564
Db 1746 GTCCTATCTGGTATGGTTAC---AATGAGAGCTAAAGCTTCGAGAGACTGGCATACA 1802
QY 2565 TCAACACCAACCATACCCCTTACCTCGCTCCCGCTGCTCGCTTACTGCACCCCTCCCG 2624
Db 1803 TCAACACCAATTGTTTATCCAATTACATCTATCCACTAGTAGCATACTGGCTCTCTG 1862
QY 2625 CGGTCTGCTCTTCAACGGCAAGTTATCATGCGCTCGATAGCAAGTTGCGCAAGCTCT 2684
Db 1863 CTATCTGTTTACTCAACCAAAATTTATTTCTCTCGATAGCAATTTATGCTGGGGGT 1922
QY 2685 TCTTCATCGCCTCTTTCATGCTCATCTTCGGAACGGGATCTCTGAGATGCGGTGAGG 2744
Db 1923 TCTTCATCTGCTTTTGTCTTCCATCTTCGCACCTGGTATTTTGGAGCTTCGATGGAGT 1982
QY 2745 GGGTGAGCATCAGGAGTGGTGGAGAAACGAGCAGTTCTGGGTCTATCGCGCGGTGTCG 2804
Db 1983 GTGTTGGCATTGAGGATGGTGGAGAAATGAGCAGTTTGGGTCAATTGGTGCACTCTG 2042
QY 2805 CGCATCTCTTCGCGCTGTCGAGGCGCTGCTCAAGGCTCTCGCGGATTCGACACCAACT 2864
Db 2043 CACATCTCTTGTCTGTGTTTCAAGGCTCTTTAAAGTGTCTAGCAGGGATTCGACACAAACT 2102
QY 2865 TCACGCTCACCTTCAAGGCCACCGGCGAGGACGAGGTTTCGCGAGCTCTACGCT 2924
Db 2103 TCACGCTCACATCAAGGGAAC---CGATGATGATGGTGAATTTTGTCTGAGCTGTATGT 2159
QY 2925 TCAAGTGAACCAAGCTCTCTATCCCGCCCAACACGCTCTCATCAATTAACGTCATCGCGG 2984
Db 2160 TCAAGTGAACCAACTCTTCTGATCCCCCCCCACACATCTGCTGTGATTAACCTGGTTGTA 2219
QY 2985 TGTGGCGGCACTCTCGACGCCATCAACACGGGTAACAGTCTGGGGGCGCCCTCTTCG 3044
Db 2220 TAGTGGCTGGAGTGTGATGTATCAACAGTGGCTACCAATCATGGGTCACACTATTTCG 2279
QY 3045 GCAGCTCTTCTCGCTTCTGGGTTCATGCTCCACTCTACCGGTTCTCTCAAGGGGCTCA 3104

Db 2280 GGAAGCTGTTCTTTGCAATCTGGGTGATCTCCACCTCTACCTTTCTGGAAGGTCTCA 2339
QY 3105 TGGGCGCCAGAACAGGACGCCACCGTTGTGTGTCATCTGTCATCTGCTGGCTCCA 3164
Db 2340 TGGGAAGCAGAACCGCACACCGACCATCGTCAATGTTGTCCTCTCTTCTGCTCCA 2399
QY 3165 TCTTCTCCCTGCTCTGGGTGATCGAGTCCAGCCCTTTTCATCTGTCAGGACC---AAGGCGCCGG 3221
Db 2400 TATTCTCGCTCTGCTGGGTGAAGATCGACCCCTTCATATCCCTACCCAGAAGGCTCTTT 2459
QY 3222 ACGTCAGGACGATGTGGCATCAATGTCGAGCTGTT 3256
Db 2460 CCCGTGGGACGTGGTGTAAACTGCTGAATGAT 2494

RESULT 12
CL964957
LOCUS OsIFC0011369 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
CL964957
ACCESSION CL964957.1 GI:52384602
VERSION
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 2595)
AUTHORS Ma,L., Wang,C., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Chinese Academy of Sciences
Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source location/Qualifiers
1..2595
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Query Match 23.0%; Score 791.2; DB 9; Length 2595;
Best Local Similarity 61.7%; Pred. No. 5.8e-146;
Matches 1546; Conservative 0; Mismatches 813; Indels 147; Gaps 12;
QY 834 AGCCGCTGTCAGGAAGGTGTCGATCGCTCGAGCAAGGTGAACCGTACCGGATGGTGA 893
Db 92 AGCTCTCAGCAGAGGCTGTCCGTCGCTCGGGAGCTCAACATGTACCGCGCGCG 151
QY 894 TCGTGGTGGTCTGCTGTGTGCTCGCTTCTTCTCCGATACCGTATCTCTGACCCCGGCC 953
Db 152 TCGGCTCGGCTCGTCTCATCGCGCTTCTTTCGCTACCGGTACCCCGCGCTCG 211
QY 954 CGGAGCGCATCGGCTGTGGCTCTCTCCATCATCTGGAGATCTGGTTCGCACTCTCT 1013
Db 212 CCGAGCCCAACGCGCTGTGGGTGACGCTCCGCTCGAGCTCTGGCTCGCGCTCAT 271
QY 1014 GGATCTCGACAGTTCGCCAAGTGGTTCCCATCTGACCGGAGACGTACCTCGACCGCC 1073
Db 272 GGCTGATCGCCCACTCCGGAAGCTCTCCCGCGCAACCGGCTCAGTACTCGACAGGC 331


```

Db      2465  ACCGGGCGCGGACCATCGCGTCTCTGGTGGTGGTCTTCTGGTCTCCGTGGGTGCG 2524
QY      3174  TGCTCTGGGTGAGATCGACCTTTTCATCGTCAGGACCAAGGCC 3219
Db      2525  TGCTCTGGGTGAGATCGACCTTTTCATCGTCAGGACCAAGGCC 2570

RESULT 13
CD726831
LOCUS   1874 bp mRNA linear EST 26-JUN-2003
DEFINITION Cucurbita pepo testa subtracted cDNA Cucurbita pepo cDNA
clone CES similar to Cellulose synthase, mRNA sequence.
ACCESSION CD726831
VERSION   1
KEYWORDS EST.
SOURCE   CD726831.1 GI:32277678
ORGANISM Cucurbita pepo
           Cucurbita pepo
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
           1 (bases 1 to 1874)
REFERENCE Bezdol,T.N., Mathews,D., Loy,J.B. and Minocha,S.C.
AUTHORS Molecular analysis of the hull-less seed trait in pumpkin:
TITLE Expression profiles of cell wall related genes during development
JOURNAL Unpublished (2003)
COMMENT Contact: Subhash Minocha, Dennis Mathews, Brent Loy, Todd Bezdol
Dr. Minocha
University of New Hampshire
Rudman Hall, Durham, NH 03824, USA
Tel: 603 862 3840
Fax: 603 862 3784
Email: sminocha@cisunix.unh.edu
Degenrate primers and Taq were used to amplify cDNA for TOPO TA
(Invitrogen, Carlsbad, CA) cloning. Sequencing was performed three
times using the Dynamic ET Terminator Sequencing kit (Amersham
Pharmacia Biotech Inc, Piscataway, NJ). Sequences were identified
by NCBI BLAST(X).
FEATURES             Location/Qualifiers
     source           1..1874
     /organism="Cucurbita pepo"
     /mol_type="mRNA"
     /db_xref="taxon:3663"
     /clone="CES"
     /dev_stage="20 days post-anthesis"
     /clone_lib="Cucurbita pepo testa subtracted cDNA"
     /note="Organ: Testa; Total RNA was isolated from 20 day
post-anthesis testa tissue and used in a subtraction
hybridization procedure as according to the Clontech
PCR-Select cDNA Subtraction kit (PT1117-1) (Clontech, Palo
Alto, CA)."
ORIGIN
Query Match      22.7%; Score 782.4; DB 6; Length 1874;
Best Local Similarity 64.7%; Pred. No. 3.1e-144;
Matches 1213; Conservative 0; Mismatches 601; Indels 60; Gaps 1;
QY      1015  GATCTCGACCAAGTTCCCAAGTGGTTCCCATCGACCGGAGAGGTACTCGACCGCT 1074
Db      1    GGTGTGGACCAAGTTCCCAAGTGGTTCCCATCGACCGGAGAGGTACTCGACGAT 60
QY      1075  CTCCTCAGGTACGAGAGGAGGAGCGGTGCTGCTGTCGGGGTGGACCTTTCGT 1134
Db      61  ATCAGCCAGGTTTGAAGAGAGGAGACCTTCCCAAGTGGTTCCCAAGTGGTTTGT 120
QY      1135  GAGCAGGTGGACCGCTCAAGAGCGCGGTGGTGGACCGCAACCGTCTCCAT 1194
Db      121  GAGTACCGTGTATCGGTGAAGAAGAACCTCTCTAATCACTGGGAATACCGTCTTCAAT 180
QY      1195  CCTCGCGGTAGACTACCGCGTGGACAAGGTCTCTGCTACGTCTCCGACGAGCGGCGTC 1254
Db      181  CCTGTGCTGGACTATCCTGTGATTAAGTCTCTGCTATGTCGGATGACGGTGCAGC 240
QY      1255  GATGCTGAGTTGCGATGCTCTCGGAGACGCGGAGTTCCGCGCAAGTGGTGGCCCTT 1314

```

```

Db      241  TATGCTCACAATTGAATCTCTAGTTGAAACAGCTGACTTTGCAAGGAATGGGTTCCT 300
QY      1315  CTGCAAGAAGTTGGGCATCGAGCCCGCGCCCGGAGTTCTACTTCTCGCTCAAGGTGCA 1374
Db      301  CTGCAAAAAAATCTCCATTGAACCTCGAGCTCTCTGAGTTTACTTCTCCAGAAATAGA 360
QY      1375  CTACCTCAAGGACCAAGGTGCGAGCCCACTTCTGTCAGAGCGCCGCGCATGAAGAGAGA 1434
Db      361  CTATTGGAAGATAAAGTACAACTCTTTTGTAAAGCAAGCTAGAGCTATGAAGAGAGC 420
QY      1435  GTATGAGCAGTTCAAGTCCCGATCAACGCGCTGGTGGCCAGGSCCATGAAGTCCCGC 1494
Db      421  TTATGAAGAGTTTAAAGTACGAGTCAATGCTTTGGTGGCAAGGCGCAAAAAACCTTGA 480
QY      1495  AGAGGGGTGGATCATGAAGAGGACGCGACCGCTGGCCCGGGAACAAACACCGCGCACCC 1554
Db      481  TGAAGGCTGGTCCATGCAAGATGGAACAGCTTGGCCAGGAACAACATCCACGTGATCAC 540
QY      1555  CGCATATGATCCAGGTGTTCTTGGGGCCACAGCGGGCGCCACGACACCGAGGGCAACGAG 1614
Db      541  TGGAAATGATTCAGGTTTCTTGGAAAATACCGGGCGCCCATGACGTAGAGGGAAATGAACT 600
QY      1615  GCCCGCCTCTGTCAGCTCTCCGTCGAGAAGGCGCGGGATTCAGACGACCAAGAGAGC 1674
Db      601  TCCTCGACTAGTCTACGCTCTCAGAGAGAGAGGCGCTGTTACCGAGCATCAAAAAAGC 660
QY      1675  CGGCGCCATGAACGCTCTGATTGCGTCTCCGCGTGTCCGACCAACCGGCCATTCATGCT 1734
Db      661  TGGTGTGCAAAATGCTCTGGTAAGATATCTGCAGTCTCTCAAAATGCCCCCTTCATTT 720
QY      1735  CAACTTGGACTGTGATCACTATCAACAAACAGCAAGGCCATCCGGGAGGCCATGTGCTT 1794
Db      721  TAACTTGAATGATGATCACTATTAACAAATAGTCAAGCTATACGTAGAGGCAATGTGTT 780
QY      1795  CCTCATGACCCCTCAGTCTCGCGGAGGTCTGCTACGTTTCAGTTCGCGAGAGGTTGCA 1854
Db      781  CTTGATGAGACCCACAAGTTGGCGGAGATGTAATGTATGTTTCTCTCAGAGGTTTGA 840
QY      1855  CGGATTCGATCGACGACCGATACGCTTAAACAGGAACACCGCTCTTCTTCGACATCAACAT 1914
Db      841  TGGCAATTGATCGAAGTGTATGATGCAATGCAATCGCAACACGGTTTCTTCGATGTGAACAT 900
QY      1915  GAAAGGGTGTGACCGCATTCGAAGCGCCGGTGTACGTGGGACAGGGTTCGTTCCGGCG 1974
Db      901  GAAAGGATTTGATGGCATCCAGGACCAAGTTTATGTTGTTGTTGTTGTTTCAATAG 960
QY      1975  CCAGGCGTCTACGGCTACACCTTCCCAAGGGACCCCAAGAGGCCCAAGATGGTCACTG 2034
Db      961  ACAGGCACTTATGGCTACGGACCTCTTCCATGCTAGCTTATCTAAGACATCTCTCATC 1020
QY      2035  CGACTGTCGCCGTGCTTCGGCGCGCAAGAGCGGAACACACCGCAAGGACGGCTGCCGA 2094
Db      1021  ATCTTGCTCTTGTTGGTGGCTGCTGCTCTTGTGCTGCTGCTTCCCAAGAAAGATCTCAAAAGA 1080
QY      2095  GGGCAACCGCTGATATGGGAGTAGATAGCGACAAGAG-----ATGCTCATGTCCTCAATGAACTT 2131
Db      1081  TCCGACTGAGATTGAGAGATGCAAAAAGAGAGAGCTTGATGCTGCAATCTTTTAACT 1140
QY      2132  -----ATGCTCATGTCCTCAATGAACTT 2154
Db      1141  CAGGGAATAGATAATTATGATGATGATGAGAGATCAATGCTGATTTCTCAACTGAGCTT 1200
QY      2155  CGAGAAGCGGTTCGGGACGTCCGGCGCTTCCGTACGTCGACGCTGATGGAGAGGAGCGG 2214
Db      1201  TGAGAAAACCTTCGGATTGCTCTGTTGTTTATGAGTCTACGCTAATGGAATAGGCGG 1260
QY      2215  CGTCCCTCTCTCGTCGAGCCCGCGCGCTCCTCAAGGAGGCCATCCATGTGTCAGCTG 2274
Db      1261  AGTTGGGAAATCTGCAAAATCCCTCGACTTTTGTATCAAGGAAGCAATTCATGTCTAGCTG 1320
QY      2275  CGGCTACGAGGACAAAGACCGGCTGGGGCTGGAGTGGGTTGATCTACGGGTGATCAG 2334

```

1321	Db	TGGTTATTGAAGAGAAGTCCACTTGGGGGAAAAGAGATTGGTTGGATATATATGGTTCAGTGCAC	1380
2335	Qy	GGAGGACATCTCTGACGGGGTTCAAGATGCATCTGCCGGGGTGGCGCTCCGTGTACTGCAT	2394
1381	Db	TGAGGATATCTTAACTGGGTTTCAAGATGCATTCGCGAGGGTGGAGGTCCATCTACTGCAT	1440
2395	Qy	GCGGAAGCGGGGGGTTCAAGGGGTGGCGCGGATCAATCTATCGAGCGTCTCAACCA	2454
1441	Db	GCCATTTAGGCCAGCAATTCAAAGGGTCCGCACCAATTAACCTTTCTGATCGTCTCCACCA	1500
2455	Qy	GTGTCTCGGTGGCGCTGGGGTCCGTGAGATCTTCTTACGCGGCACAGCCCGCTGCT	2514
1501	Db	AGTTCTTCGATGGGCATTTGGATCTGTGTGAGATTTTCTTAGCAGACACTGTCCATATG	1560
2515	Qy	GTACGGCTACAAGAACGGCAACTCAAGTGGCTGGAGCGCTTCGCCCTACATCAACACCAC	2574
1561	Db	GTATGGATTTGACGGGGCGGCTCAATATGGCTCCAAGATGGCTTACATAAACACCAT	1620
2575	Qy	CATCTACCCCTTCACTCGCTCCCGCTCGCTGCTTACGTACGTACCAACCTCCCGCGCTGCT	2634
1621	Db	TGTCTATCCCTTCACTCGCTCCCTCTTGTGTGCTTACTGTCTCATTTGCTGCAATCTGCCT	1680
2635	Qy	CCTCAGCGCAGTTTCATCATGCTCGTTCGATTAGCAGGTTCGCCAGCTCTTCTTCATCGC	2694
1681	Db	GCTCACAGGAAAGTTTCATCATTTCCAAACGCTCTCGAACCTAGCAAGTACCCCTTTTCTCG	1740
2695	Qy	CCTCTTCATGTCATCTTCGACGGGATCTCTGAGATGCGGTGGAGCGGGTGAGCAT	2754
1741	Db	TCTCTTGTTCATCTCTCACAGTGTCTCTGAGCTGCTTGGAGTGGTGTGATGCAT	1800
2755	Qy	CGAGGATGGTGGAGAAACGACAGTCTTGGGTTCATCGCGCGGTGTCGCGCATCTCTT	2814
1801	Db	CGAAGATATATGGGTAAACGAGCAATTTCTGGGTAACTGGAGCGCTCCCGGCACATCTCTT	1860
2815	Qy	CGCGCTGTGCAGG 2828	
1861	Db	TGCGGTCTTCCAAG 1874	
RESULT 14			
CF430024			
LOCUS		789 bp mRNA linear	EST 03-SEP-2003
DEFINITION		PHI_25_C04_g1_A002 Phosphorous-deficient seedlings Sorghum bicolor	
ACCESSION		cDNA clone PH1_25_C04_A002 5', mRNA sequence.	
VERSION		CF430024	
KEYWORDS		EST.	
SOURCE		CF430024.1 GI:34442725	
ORGANISM		Sorghum bicolor (sorghum)	
REFERENCE			
AUTHORS		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.	
TITLE		1 (bases 1 to 789); Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C., Sun,F., Sullivan,R., Harris,K., Eastman,A. and Pratt,L.H.	
JOURNAL		An EST database from Sorghum: phosphorous-deficient seedlings	
COMMENT		Unpublished (2003)	
		Other_ESTs: PH1_25_C04.bl_A002	
		Contact: Cordonnier-Pratt MM	
		Laboratory for Genomics and Bioinformatics	
		The University of Georgia, Department of Plant Biology	
		Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA	
		Tel: 706 542 1860	
		Fax: 706 583 0210	
		Email: mmpratt@uga.edu	
		Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.	

QY 2918 TAGCGCTTCAAGTGACACGCTCTCATCCGCCGCCACACGCTGCTCATCATTTAACGTC 2977
|||||
Db 721 TAGCGATTCAAGTGACACGCTCTCATCCGCCGCCACACGCTGCTCATCATCAACATC 780
|||||
QY 2978 ATCGGCGTC 2986
|||||
Db 781 ATCGGCGTC 789
|||||

RESULT 15
CG454544/c
LOCUS
DEFINITION PUIKA96TDB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBta0601023,
genomic survey sequence.
ACCESSION CG454544
VERSION CG454544.1 GI:34839544
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 943)

REFERENCE Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reanick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUIKA96TDB PUIKA96TBB
CONTACT: Cathy Whitelaw
TIGR

7912 Medical Center Drive, Rockville, MD 20850, USA
Tel.: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: Sheared ends.

FEATURES
source location/Qualifiers
1..943
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBta0601023"
/clone_lib="ZM 0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

ORIGIN

Query Match 21.7%; Score 747.6; DB 9; Length 943;
Best Local Similarity 96.9%; Pred. No. 2.2e-137;
Matches 762; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 2657 CCGTCGATTAGCAGCTTCGCAGGCTCTTCTTCATCGCCCTCTTCATGTCCTTCGCG 2716
Db 908 CCGAGATTAGCAGCTTCGCAGGCTCTTCTTCATCGCCCTCTTCATGTCCTTCGCG 849
QY 2717 ACGGCATCTCGGAGATCGGTGGAGCGGGTGGAGATCGAGGAGTGGTGGAGAACGAG 2776
Db 848 ACGGCATCTCGGAGATCGGTGGAGCGGGTGGAGATCGAGGAGTGGTGGAGAACGAG 789
QY 2777 CAGTTCGGGTTCATCGCGGGGTGTCGGCGATCTTTCGGCGTGGTGGAGGGCTGCTC 2836
Db 788 CAGTTCGGGTTCATCGCGGGGTGTCGGCGATCTTTCGGCGTGGTGGAGGGCTGCTC 729
QY 2837 AAGTCTCTCGCGGGATCGACCAACTTACCGTCACTCAAGGCCACCGCGGACGAG 2896
Db 728 AAGTCTCTCGCGGGATCGACCAACTTACCGTCACTCAAGGCCACCGCGGACGAG 669
QY 2897 GACGACGAGTTTCGCGAGCTCTAGCGCTTCAAGTGGACCGCTCTCATCCGCCACC 2956
Db 668 GACGACGAGTTTCGCGAGCTCTAGCGCTTCAAGTGGACCGCTCTCATCCGCCACC 609

QY 2957 ACGCTGCTCATCATTTAAGTTCATCGGCGTCTGCGCGGCGCATCTCCGACGCCATCAACAAC 3016
Db 608 ACGCTGCTCATCATTTAAGTTCATCGGCGTCTGCGCGGCGCATCTCCGACGCCATCAACAAC 549
QY 3017 GGGTACCAAGTCTCGGGGCCCTCTTTCGGCAAGCTCTTTCGGCTTCTGGGTCAATCGTC 3076
Db 548 GGGTACCAAGTCTCGGGGCCCTCTTTCGGCAAGCTCTTTCGGCTTCTGGGTCAATCGTC 489
QY 3077 CACCTCTACCGGTTCTCTCAAGGGGCTCATGGGGGCCAGAAACAGGACGCCACCGTTGTT 3136
Db 488 CACCTCTACCGGTTCTCTCAAGGGGCTCATGGGGGCCAGAAACAGGACGCCACCGTTGTT 429
QY 3137 GTCATCTGGTCCATTTCTGCTGGCCTCCATCTTCTCCCTGCTCTGGGTCAAGATCGACCT 3196
Db 428 GTCATCTGGTCCATTTCTGCTGGCCTCCATCTTCTCCCTGCTCTGGGTCAAGATCGACCT 369
QY 3197 TTCATCGTCAAGGACCAAGGGGCCCGAGCTCAGGCAGTGTGGCATCAATTTGCTGAGCTGT 3256
Db 368 TTCATCGTCAAGGACCAAGGGGCCCGAGCTCAGGCAGTGTGGCATCAATTTGCTGAGCTGT 309
QY 3257 TATTAAGGTTCAAAATTTCTGGAGCTTGTGCATAGGGAGAAAAACAATTTAGAAATTTT 3316
Db 308 TATTAAGGTTCAAAATTTCTGGAGCTTGTGCATAGGGAGAAAAACAATTTAGAAATTTT 249
QY 3317 GTAAGGTTGTGTCGTCTGTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTT 3376
Db 248 GTAAGGTTGTGTCGTCTGTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTT 189
QY 3377 AAGGACAAAGGTTTTCATTTTAAATGGCAAAAAAAGGACGAGGAGGAGGAGGAGGAGG 3436
Db 188 AAGGACAAAGGTTTTCATTTTAAATGGCAAAAAAAGGACGAGGAGGAGGAGGAGGAGG 129
QY 3437 AAAAAA 3442
Db 128 TAGCAA 123

Search completed: March 11, 2005, 23:55:22
Job time : 7046 secs

This Page Blank (uspto)

CC ABA67073 to ABA67907, their (reverse) complements, sequences producing an
 CC Expectation (E) value of 0.01 or less compared to the 835 sequences
 CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
 CC 835 sequences or sequences that are degenerately equivalent or allelic to
 CC the 835 sequences. The polynucleotides are used to modify the activity of
 CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
 CC plant. They are especially used to modulate or alter the polysaccharide
 CC content, composition or structure of the plant. ABA16268 to ABA16340 are
 CC proteins encoded by some of the polynucleotide sequence given in the
 CC present invention
 XX
 SQ Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3851
 Score: 4470.00 Matches: 838
 Percent Similarity: 84.47% Conservative: 92
 Best Local Similarity: 76.11% Mismatches: 105
 Query Match: 79.14% Indels: 66
 DB: 3 Gaps: 16

US-10-627-132-30 (1-1052) x ABA67114 (1-3851)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
 DB 153 ATGGAACCGACGCGCGCTGGTTCGCGGTTCATACAGAAACGAGTTCTGGTGCATC 212
 QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluLeuCys 40
 DB 213 CATGACATGAGGAGCGAAGCCTTTGAACACGCTTGAGTGGCCACGCTGCCGAGTTGT 272
 QY 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
 DB 273 GGCAGACGTCGCGCGCTTAACACAGACGCGGAGCTGTTCGTTCCTGTAATGAGTGGCG 332
 QY 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
 DB 333 TTTCCTCTGTCTGCGCGCTGCTATGAGTACGAGACGAGAGAGAGAAATCAGTCGTGCCCG 392
 QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
 DB 393 CAGTGCAATCTCGTTACACGCTCAAAAGAGGAGTCCACGCGTGGAAAGGTGACAGATGAT 452
 QY 101 GluGluAspLeuAspLeuGluHisGluPheAsnLeuAspAspGluAsnGlnArg 120
 DB 453 GAAGACAGCTGTGATGACATGAACATGAATTAATGTGGAGACTCAGCAAGAAACAGG 512
 QY 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnLeuThrGluAlaMetLeuHisGlyArgMet 140
 DB 513 CAG-----CAGATCACCGAGGCGATGCTCCACGAGCGCATG 548
 QY 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsn-----AsnThrProGlnLeuPro 158
 DB 549 AGCTATGCCGAGGTCCCGACGAGGAAATTCGACAGATTGTCTATAATTCACAGAGCTTCCT 608
 QY 159 -----ProIleThrGlySerArgSerValProValSerGlyGluPheProIle 175
 DB 609 CCGCAGATTCTGTACTTGCAACGCGCCACTCGGT-----GTGAGTGGGAGATTCCA--- 662
 QY 176 ThrAsnGlyTyrGlyHisGlyGluVal-----SerSerSerLeuHisLysArgIleHis 193
 DB 663 ACGTACATACTACGACACACCAATTCCTGCCAACCCCTGCAATCTCTGAAGCGGTGTCAT 722
 QY 194 ProTyrProValSerGluProGlySerAlaLys-----TrpAspGluLysGlu--- 210
 DB 723 -----CCAAGCTCCGAGCGCGGAGTGGAAAGGATCATCATGATCCAAACAGAGGATATT 776
 QY 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLys 223
 DB 777 GGTCTTATGCTTTGGGAACGTGCTTGGAAAGGAGCGGCGATGTTATTAATTCGAAG 836
 QY 224 GlnGlyIleLeuGly-----GlyGly--- 230

DB 837 GAAAAAATAATCAGGCCAGTTGGATATGACGGAAGGGAGATATCAATATATATGGGGGTTT 896
 QY 231 -----AlaAspProGluAsp---MetAspAlaAspValAlaLeuAsnAspGluAlaArg 247
 DB 897 GCACCAAAATGAGCGTGAAGATTATATTGATCCCGATATGCCAATGACCCGATGAAGCAGG 956
 QY 248 GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetVal 267
 DB 957 CAGCCACTGTCCCGAAAGTGCCTTCCTCAAGCAAAATAAATCCATCAGCAATGGTC 1016
 QY 268 IleValValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProVal 287
 DB 1017 ATTGTAATTCGATGATGCTGGGTATTTTCTCCGCTATGCTCTCTCCATCCAGTG 1076
 QY 288 ProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSer 307
 DB 1077 AAGATGCAATATGGGCTCTGGGCCACTTCTATGCTTTGTGAAATCTGGTTTGGCTGTGTC 1136
 QY 308 TrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArg 327
 DB 1137 TGGATTCTTGATCAGTTTCCCAAGTGGTTCCTATCAGTCGTGAAACGCTATCTCGATCGA 1196
 QY 328 LeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe 347
 DB 1197 CTGTCAATTAAGGTACGACGAGAGCGCAACCATCAATGCTTGCCTGTGTGACCTCTTT 1256
 QY 348 ValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSer 367
 DB 1257 GTGAGTACTGTAGATCCACTGAAGGAGCCTCCTTGTGTTACTGCCAATACAGTATTATCA 1316
 QY 368 IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAla 387
 DB 1317 ATCTTTTCAGTAGACTACCTGTAGACAATGCTCTCTGTTATGCTCTCTGATGCGGAGCG 1376
 QY 388 SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro 407
 DB 1377 TCGATGCTTACTTTTGAATCTCTCTGAGACCTCAGAAATTTGCCAGAAATCGGTACCA 1436
 QY 408 PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysVal 427
 DB 1437 TCTCGAAGAAATTCGACATTGAGCTTCGCGCTCCCGAAATCTATTCTCTCAGAAATTT 1496
 QY 428 AspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg 447
 DB 1497 GACTATCTGAAGGACAAATTTCAACCCACCTTTGTCAAAGAGCGCGTGCATGAAGAGA 1556
 QY 448 GluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValPro 467
 DB 1557 GAATATGAAGAATTCAGGTGCGCATCAATTCGGTGTGTTGCAAGGCGCTCTAAAGTGCCC 1616
 QY 468 AlaGluGlyTrpIleMetLysAspGlyThrProTyrProGlyAsnAsnThrArgAspHis 487
 DB 1617 AAGAGAGGATGGCAATTCGAAGACGTGACGCTTGGCTGGTAAATAATACCCGTGACCAT 1676
 QY 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGlu 507
 DB 1677 CTTGGTATGATCCAAAGTGTCTTGGGTCAAGTGGCGGCTCGATACAGAAGGCAATGAG 1736
 QY 508 LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLys 527
 DB 1737 CTTCTCCGGCTAGTATATGTTTCTCGTGAAGAGACCTGGTTTCCAGCATCACAGAAG 1796
 QY 528 AlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMet 547
 DB 1797 GCCGGTCCCATGAATGTTTGGTTCGGGTTTCTGCTGTCTCACCATGCTCCATTATG 1856
 QY 548 LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
 DB 1857 CTGAATCTGATGATGATCACTACATTAACTAAGCAAGCAATCAGGAAGGCGATGTGC 1916
 QY 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPhe 587
 DB 1917 TTTATGATGATCCTCAGGTGGGAGAAAGTCTGTTATGTCCAATTTCCCTCAGAGATTC 1976

QY 588 AspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsn 607
DB 1977 GATGGTATTGATCGCAATGACCGTTACGCCAATCGAAACACCGCTATTCTTGTATATCAAC 2036
QY 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArg 627
DB 2037 ATGAAGGCTCGATGGAATTCAGGGCTGTATATGTGGAACTGGATGCAATGTCAGA 2096
QY 628 ArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThr 647
DB 2097 AGACAAGCTCTATATGGGTATGGGCTCCCAAGGCCCAAAAGCTCCCAAGATGTGACC 2156
QY 648 CysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLys----- 663
DB 2157 TGTGATTGTCTCCCTTGTGTGGGTCTCTCGTAAGAGTCTCCGAAGAAAAATAGTAGCAAG 2216
QY 664 -----AspGlyLeuProGluGlyThrAla 671
DB 2217 AAAAGTCAGGAATCCAGCTCCCGCTACAACTCGACGGGATCAGGAAGAGTAGAA 2276
QY 672 AspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArg 691
DB 2277 -----GGTTATGATGACGAAGAGCAATTGTTGATGAGCAACTAGACTTCGAGAAGAG 2330
QY 692 PheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyGlyValProPro 711
DB 2331 TTTGCCAGCTCTTACGCTTTTGTCAATCCACTCTGATGGAGATGGTGTGTTCCGCA 2390
QY 712 SerSerProAlaAlaLeuLysGluAlaIleHisValIleSerCysGlyTyrGlu 731
DB 2391 ACAGCAATCCAGCTGAATTGTTGAAGGAGGCTATTTCATGTCATCAGCTGTGGATATGAA 2450
QY 732 AspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIle 751
DB 2451 GACAAACCGAATGGGAAAGAGCTTGGATGATCTATGGATCAGTCACAGAGCAATT 2510
QY 752 LeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArg 771
DB 2511 CTGACTGGATTCAAGATGCACACTCGAGGCTGGCGTCCATTACTGTATGCCAAGCA 2570
QY 772 AlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg 791
DB 2571 CGACATTTCAAGGGTCTGCTCCAACTCAATCTATCAGACCGGTTGAACGAGTGTGGT 2630
QY 792 TrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyr 811
DB 2631 TGGGCTTTGGGATCAGTAGAATTTTCATGACGACACATTCGCCAATCTGGTATGGCTAT 2690
QY 812 LysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 831
DB 2691 ---GGGGGAGGCTCGAAATGGCTTGAAAGATTGGCTATATCAACACCATTTGCTATCCA 2747
QY 832 PheThrSerLeuProLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGly 851
DB 2748 TTCACCTCTTCCACTCATTCGCTATTGACACTTCCAGCGCTCAGTTTGTCTACTGCG 2807
QY 852 LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMet 871
DB 2808 AAAATTGGATCCCTCAGATCAGTACTTTTGCAGTCTATTTTATAGTCTTTTTCATC 2867
QY 872 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluTrp 891
DB 2868 TCAATTTTTGGCACTGTATTCTGAATAGGTTGAGTGGAGTGGAGTGGAGTGGAGTGG 2927
QY 892 TrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValVal 911
DB 2928 TGGCGAAATGAACAGTCTCGGTTATTCGAGGGGTTCTCGACATTTTTCGAGTATT 2987
QY 912 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 931
DB 2988 CAAGGTCTGCTCAGGTTACTGGCAGGCAATTGATACAAATTTTCACAGTCACTGCCAAGCA 3047

QY 932 ThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeu 951
DB 3048 TCA-----GATGACGGTGAGTTTGGGGAACCTGTATGCAATCAAATGACACACTCTC 3101
QY 952 IleProProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAsp 971
DB 3102 ATTCTCTCTACAACTCTGTGTCATCAACCTTGTGGGGGTGGTGTGGGGGTAGCAGAT 3161
QY 972 AlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 991
DB 3162 GCAATCAACAATGGATTTTCAGTCATCGGGTCTCTCTTGGGTAAAGCTTTTCTTTCATTC 3221
QY 992 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
DB 3222 TGGGTCAATTGTCACCTGTATCTTCTCAAGGGTCTCATGGGCGAGCAACCGAACA 3281
QY 1012 ProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1031
DB 3282 CCCACCTCGGTGTTATTTGGTCAATCTGCTGGCATCTGTTTCTCTTTCTTCTGGGTA 3341
QY 1032 ArgIleAspProPheIleValIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn 1051
DB 3342 AGAATTGATCTTCTTGGTAAAGTTAAAGGCCAGATACTAAACAATGTGGCATCAAC 3401
QY 1052 Cys 1052
DB 3402 TGC 3404
RESULT 3
AAC83247
ID AAC83247 standard; DNA; 7234 BP.
XX AAC83247;
XX 14-MAR-2001 (first entry)
XX Cellulose synthase polynucleotide sequence SEQ ID 1.
DE Cellulose synthase; lignin; secondary cell wall construction; wood pulp;
KW transgenic plant; paper manufacture; ds.
XX Arabidopsis thaliana.
OS WO200070058-A2.
PN 23-NOV-2000.
XX 17-MAY-2000; 2000WO-GB001890.
XX 18-MAY-1999; 99GB-00011379.
XX (UTMA-) UNIV VICTORIA MANCHESTER.
XX Turner S, Taylor N;
XX WPI; 2001-041015/05.
DR Cellulose synthase gene expressed during deposition of secondary cell
PT walls in lignin-containing cells, useful for modulating expression of
PT enzymes involved in synthesis of plant cell walls and to produce
PT transgenic plants.
PS Claim 3; Page 28-30; 49pp; English.
XX This invention relates to a cellulose synthase gene expressed during the
CC deposition of secondary cell walls in cells containing lignin. The
CC cellulose synthase gene is useful for regulating the expression of genes
CC specifically during secondary cell deposition in lignin containing cells.
CC It can be used to modify the structure and cellulose content of plant
CC secondary cell walls and to produce altered plant phenotypes specific
CC to the needs of a particular industry such as in reducing the lignin of wood
CC pulp for paper manufacturing. A construct containing a cellulose synthase
CC promoter sequence and a gene of interest may be used in a method for the

CC production of the product of the gene of interest in a host cell that
CC produces lignin, where the product is produced only during secondary cell
CC wall synthesis. The present sequence represents the *Arabidopsis thaliana*
CC cellulose synthase gene of the invention

Alignment Scores:	
Pred. No.:	0
Score:	424.50
Score Similarity:	61.35%
Best Local Similarity:	55.69%
Query Match:	74.80%
DB:	4
Length:	7234
Matches:	856
Conservative:	87
Mismatches:	90
Indels:	508
Gaps:	20

US-10-627-132-30 (1-1052) x AAC83247 (1-7234)

Qy	1	MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle	20
Db	92	ATGGAAGCTAGCGCGGCTTTGTGCGCGGTTCTCATAAACCGTAATGAACCTAGTCGTCA	151
Qy	21	ArgGlyHisGluAsp-----	25
Db	152	CACAACCAATGAAGAGTGTTCACATTTACTTTTTTCTCATCTACCTTACCAGAGTTTTTTT	211
Qy	25	-----	25
Db	212	TACCACGCTAGTAAAAATATTATTCGATTTTTTCGTTTTATTGTGTTACTATTCTTAA	271
Qy	25	-----	25
Db	272	AATTCTGATGTTTTGGGAATAAAATATGATCATTTTTTTTAAATCATCTTCTTATTATGA	331
Qy	25	-----	25
Db	332	GACAAAATTTATATCTGTATTCTGTAGTTGCAATAATGTTGTAGAAAATTCATATCTT	391
Qy	25	-----	25
Db	392	TGTTAGCAACATATAAATTTTTTGGTAAATTAAGTTGAGAACTCAGGTTTAAACCAT	451
Qy	25	-----	25
Db	452	TTAATCGCTGTCATTTTTTTTATATCTTTATCTTCACTTAGAATCTTTTTTTGTTGATAG	511
Qy	26	-----Pro 26	
Db	512	AAATTTTCAGGTTTTATTTCGTCTTTAAGGAACCTTAAACTTTTGTGTAATTATAACAGCCA	571
Qy	27	LysProLeuArgAlaLeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeu	46
Db	572	AAGCCTCTGAAGAACTAGATGCACAAATCTCTGTGAGATATGTGGAGATCAGATCGTTTA	631
Qy	47	ThrValAspGlyAspLeuPheValAlaCysAsnGluCysGlyPheProValCysArgPro	66
Db	632	ACAGTAGAAGGAGACCTCTTCGTAGCTTCAATGAGTGTGGTTTTCCGCGGTGAGACCT	691
Qy	67	CysTyrGluTyrGluArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyr	86
Db	692	TGCTATGAGTACGAGAGAAGAGAGAGAACACAAAACTGTCCCTCAGTGTAAAGACTCGTTAC	751
Qy	87	LysArgLeuLys-----	90
Db	752	AAGCGTCTCAGAGGTAAGTTATTTATATCTCCCTCTGCTCTGTGTTGTTTCGACGAAA	811
Qy	90	-----	90
Db	812	TGCCTCTATGAAAAATTTAAAAAGGCTGTTCCTTTTTTTTGTAGTTTGAACCTGGAGAGTAATG	871
Qy	91	-----GlySerProArgValAlaGlyAspAspGluGluAsp	103
Db	872	ATCTGTTTTTGGTTTTCTGAAGGAAGCCCAAGAGTGGAGGAGATCGAACGCAAGAAGAT	931

Qy 897 heTTPValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysv 917
 Db 4177 TCTGGGTCATTGGAGGAATCTCAGCTATCTCTTGGGTGTGTCGAAGGTCTCTCAAA 4236
 Qy 917 alLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspA 937
 Db 4237 TCTTAGCAGGCATTGACACAACTTCCCGTTCACATCAAGGCAACA-----GATGATG 4290
 Qy 937 spGluPheAlaGluLeuTyrrAlaPheLysTrpThrThrLeuLeuIleProProThrThrL 957
 Db 4291 ATGACTTTGGAGAACTTTACGCATTCAAAATGGACACACTGCTGATCCTCCCAACAAC 4350
 Qy 957 euLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyT 977
 Db 4351 TCTTAATCATAAACATTTGGCGTGTGTGAGGATCTCAGATGCCATTAACATGAT 4410
 Qy 977 YrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisL 997
 Db 4411 ATCAGTCTTGGGACCTCTATTGTGTAACCTCTTCTCTCTCTTCTCTCTCTCTCTCT 4470
 Qy 997 euTyrrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValI 1017
 Db 4471 TCTACCCATCTCTCAAGGCTGTGATGGGTAGACAGAACAGAACCAACCACTTGTGTGA 4530
 Qy 1017 leTTPSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheI 1037
 Db 4531 TTTGGTCAAGTGTATTGGCATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4590
 Qy 1037 leValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db 4591 TGCTCAAGACCAAGGACCTGACACTTCCAAGTGTGGCATCACTG 4637
 RESULT 4
 AAZ58263
 ID AAZ58263 standard; cDNA; 3776 BP.
 XX
 AC AAZ58263;
 DT 08-MAY-2000 (first entry)
 XX
 DE Corn cellulose synthase cDNA.
 XX
 KW Corn; maize; cellulose synthase; transgenic plant; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT 2..3448
 FT /*tag= a
 FT /partial
 XX
 FN WO200004166-A2.
 XX
 PD 27-JAN-2000.
 XX
 PF 13-JUL-1999; 99WO-US015871.
 XX
 PR 14-JUL-1998; 98US-0092844P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Allen SM, Fader GW, Falco SC, Kinney AJ, Lightner JE, Miao G;
 PI Rafalski JA, Thorpe CJ;
 XX
 DR WPI; 2000-182431/16.
 XX
 DR P-PSDB; AAZ58832.
 XX
 PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
 PT probes for isolating cDNAs and genes encoding homologous proteins, for
 PT producing transgenic plants.
 XX
 PS Claim 9; Page 37-38; 93pp; English.

XX The present sequence is that of a contig of cDNA clones and encodes a
 CC portion (see AAZ58832) of corn cellulose synthase (CS). The cDNA clones
 CC were isolated from corn cob, developing tassel, seedling, shoot culture
 CC and leaf tissue cDNA libraries on the basis of homology to Arabidopsis
 CC and cotton CS sequences. The invention relates to isolated nucleic acid
 CC fragments encoding plant CS and to CS polypeptides. It also relates to
 CC the construction of a chimeric gene encoding all or a portion of the CS,
 CC in sense or antisense orientation, where expression of the gene results
 CC in altered levels of the CS in transformed host cells. The host cells can
 CC be used to screen compounds for their ability to inhibit CS activity. CS
 CC nucleic acids are also useful for producing transgenic plants having
 CC altered levels of CS, and hence altered levels of fibre. CS may also
 CC serve as a target for the development of novel herbicides
 XX
 SQ Sequence 3776 BP; 931 A; 906 C; 999 G; 940 T; 0 U; 0 Other;
 Alignment Scores: Length: 3776
 Pred. No.: 0 Matches: 760
 Score: 4091.00 Conservative: 128
 Percent Similarity: 79.64% Mismatches: 137
 Best Local Similarity: 68.16% Indels: 90
 Query Match: 72.43% Gaps: 16
 DB: 3
 US-10-627-132-30 (1-1052) x AAZ58263 (1-3776)
 Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
 Db 179 ATGAGGCCAGCGCGCGCTGGTGGCGGCTCGCACAAACCGAACGAGCTCGTCTATC 238
 Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
 Db 239 CGCGCGATGGCGAGCCAGCGCGGCGAGCCCATGGACCGAGCGGAGCGGAGTGTGCAG 298
 Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
 Db 299 ATTTGCGCGACGACGTGGGGCGCAACCCGACGCGGAGCGGCTTCTGTGGCGTGTCAACGAG 358
 Qy 59 CysGlyPheProValCysArgProCysTyrrGluTrpGluArgArgGluGlyThrGlnAsn 78
 Db 359 TGGCGCTTCCCATCTCTCCGGGACTGCTACGAGTACGAGCGCGCGGAGCGGAGCGAGAAC 418
 Qy 79 CysProGlnCysLysThrArgTyrrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
 Db 419 TGCCCCCAGTGCAAGACCGCTTCAAGCGCTTCAAGGGGTGGCGCGCGTCCCGGGGAC 478
 Qy 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
 Db 479 GAGGAGGAGGACGCGCTCGACGACCTGGAGAACCGAGTTCAACTGGAGCGACAGCACGAC 538
 Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
 Db 539 TCCAGTAGCTCTC-----GCCGAGTCCATGCTCAACGCC 571
 Qy 139 ArgMetSerTyrrGlyArgGlyProAsp---AspGlyAspGlyAsnAsnThrProGlnIle 157
 Db 572 CACATGAGCTAGCGCGCGCGCGCGCGCTCGACGCGCTCGCGCGAGCCATTCCACCCCATC 631
 Qy 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
 Db 632 CCC-----AATGTTCCCTC-----CTCACCAC 655
 Qy 178 GlyTyrrGlyHisGlyGluValSerSerSerLeuHis----- 189
 Db 656 GGCAGATGTGTGATGATCATCCCGCGAGCAGCAGCGCCCTGTGCTCGTGGGT 715
 Qy 190 -----LysArgIleHisProTyrrProValSerGluProGly-----Ser 202
 Db 716 GCGCGGGGAGAGAGATTACCTCTCCCGTACGCGGATCCCAACCTTCCTGTGCAACCG 775
 Qy 203 AlalysTrpAspGluLysLysGlu-----ValSerTrpLys 214

Db 776 AGGCTATGACCCCTCCAGGATCTCGCGCATATGGCTACGGAGCGTAGCATGGAG 835
Qy 215 GluArgMetAspAspTrpIysSerIysGlnGlyIleLeu-----Gly 228
Db 836 GAGAGATGAGAGCTGGAAGCAGAGCAGGAGAGGATCCACAGACGAGGAACGATGCG 895
Qy 229 GlyGlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGln 248
Db 896 GCGCGCGATGATGGTGAT-----GATGCAGATCTACCACTAATGATGATGAAGCTAGACAG 949
Qy 249 ProLeuSerArgIysValSerIleAlaSerSerIysValAsnProIysArgMetValIle 268
Db 950 CCATTGTCAGAAAGATCCCGCTTCTTCACGCCAAATCAACCCCTATAGATGATATA 1009
Qy 269 ValValArgLeuValValAlaPhePheLeuArgIysArgIleLeuHisProValPro 288
Db 1010 ATAAATTCGGCTAGTGTTGTGTTCTTCTCCACTACCGAGTGCATCGCGTGCT 1069
Qy 289 AspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrp 308
Db 1070 GATGCATTGCTTTATGGCTCATATCTGTGATCTGTGAATTTGGTTGGCCATGCTTGG 1129
Qy 309 IleLeuAspGlnPheProIysTrpPheProIleAspArgGluThrTrpLeuAspArgLeu 328
Db 1130 ATTCTTGACCAAGTTCCTTCTATCGAGAGGAACCTATCTGACCGCTG 1189
Qy 329 SerLeuArgTrpGluArgGluGluProSerLeuLeuSerAlaValAspLeuPheVal 348
Db 1190 AGTTTAAAGTTTGACAAGAGGCGATCTTCTCAACTCGCCCTGTTGATTTCTTTGTC 1249
Qy 349 SerThrValAspProLeuIysGluProProLeuValThrAlaAsnThrValLeuSerIle 368
Db 1250 AGTACGGTGTATCCCTTGAAGGAACCTCCATTGGTCACTGCTAACTGTTCTATCTATC 1309
Qy 369 LeuAlaValAspTrpValAspIysValSerCysTrpValSerAspAspGlyAlaSer 388
Db 1310 CTTTCGGTGATATCCAGTTGATAGGTTCTCTGCTACGTTCTTGATGATGCTGCC 1369
Qy 389 MetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgIysTrpValProPhe 408
Db 1370 ATGCTGACATTTGAAGCATTTGCTGAACATCTGAAATTTGCAAGAAATGGTTCCTTTC 1429
Qy 409 CysIysIysPheGlyIleGluProArgAlaProGluPheTrpPheSerLeuIysValAsp 428
Db 1430 TGCAAAAGATATAGCCTTGAGCTCGCTCCAGAGTGCTCTTCCACAGAAAGATAGAC 1489
Qy 429 TyrLeuIysAspIysValGlnProThrPheValGlnAlaArgAlaMetIysArgGlu 448
Db 1490 TACCTGAAGACAGAGTGGCGCAAACTTTGTTAGAGACGGAGAGCAATGAGAGAG 1549
Qy 449 TyrGluGluPheIysValArgIleAlaLeuValAlaLysAlaMetIysValProAla 468
Db 1550 TATGAGGAATTCAGAGTCAGATCAATGCTTGGTTGCTTAAAGCCCAAGGTTCTCTGAG 1609
Qy 469 GluGlyTrpIleMetIysAspGlyThrProTrpProGlyIysAsnThrArgAspHisPro 488
Db 1610 GAAGATGAGCAATCAGAGATGGAATCCATGCGCCGGAATAATGTCGCGTATCATCT 1669
Qy 489 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyValGlnLeu 508
Db 1670 GGAATGATTCAGGTTTTCTTGCTCAAGTGTGGCCATGATGTGAAGGAATGAGCTG 1729
Qy 509 ProArgLeuValTyrValSerArgGluIysArgProGlyPheGlnHisIleIysIysAla 528
Db 1730 CCTCGATTGTTTATGTTTCAAGAGAAAAAGCCAGGCTACAACCATCACAGAAGGCT 1789
Qy 529 GlyAlaMetAsnAlaIleArgValSerAlaValLeuThrAsnAlaProPheMetLeu 548
Db 1790 GGTGCTATGAATGATGCTGCGAGTCTGCTGTACTAACTAATGCTCTTATTTGCTG 1849
Qy 549 AsnLeuAspCysAspHisTrpIleAsnAsnSerIysAlaIleArgGluAlaMetCysPhe 568
Db 1850 AACTTGGATTGTGATCACTATATCAATAATAGTAGGCTATATAAGGAAGCAATGTGTTT 1909

Qy 569 LeuMetAspProGlnValGlyArgIysValCysTrpValGlnPheProGlnArgPheAsp 588
Db 1910 ATGATGGATCTTCTGCTTGGAAAGAAAGTTGCTATGTGTCAGTTCTCTCAAAGATTGAT 1969
Qy 589 GlyIleAspValHisAspArgTrpAlaAsnArgAsnThrValPhePheAspIleAsnMet 608
Db 1970 GGGATTGATGCCCATGATCGATATGCTAACAGAAATGTTGTCTTTTTCGATATCAACATG 2029
Qy 609 LysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArg 628
Db 2030 AAAGTTTGGATGGTATCCAGGGCCCATTTATGTGGGTACTGGATGTGCTTTCAGAGG 2089
Qy 629 GlnAlaLeuTrpGlyTrpAsnProProIysGlyProIysArgProIysMetValThrCys 648
Db 2090 CAGGCATTATATGGCTAGCTGCCAAA---ACAAAGAACCCACCATCAAGAACTTGC 2146
Qy 649 Asp-----CysCysProCysPheGly---ArgIysIysArgIys 660
Db 2147 AACTGCTGGCCAAAGTGTGTCATTGCTGTGTTTTTGGTAAACAGGAAGCAAGAG 2206
Qy 661 HisAlaIysAspGlyLeuProGlu----- 668
Db 2207 AAGACCAAGACCTCTAAACCTTAATTTGAGAAGATAAAGAAACTTTTAAAGAAAAGNA 2266
Qy 669 -----GlyThrAlaAspMetGlyValAspSer 677
Db 2267 AATCAAGCCCTGTCATATGCTCTTGGTGAATTTGATGAAGCGCTCCAGGAGCTGAAAT 2326
Qy 678 AspIysGluMetLeuMetSerHisMetAsnPheGluIysArgPheGlyGlnSerAlaIle 697
Db 2327 GAAAGGCTGATTTGTTAAATCAACAGAAAGTTGGAAAGAAATTTGGCCAGCTCTTCAGTT 2386
Qy 698 PheValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaIle 717
Db 2387 TTTGTTGATCCACACTCTTGGAGAAATGGTGAACCTTGAAGAGTCCAGTCCAGCTTCT 2446
Qy 718 LeuLeuIysGluAlaIleHisValIleSerCysGlyTrpGluAspIysThrAspTrpGly 737
Db 2447 CTTCTGAAGGAAGCTATACATGTCATCAGTTGTGATATGAAGACAAACAGGCTGGGA 2506
Qy 738 LeuGluLeuGlyTrpIleTrpGlySerIleThrGluAspIleLeuThrGlyPheIysMet 757
Db 2507 AAAATATTTGTTGGATTTATGGATCATGACAGAAATATTTCTTACTGGGTTTAAAGATG 2566
Qy 758 HisCysArgGlyTrpArgSerValTrpCysMetProIysArgAlaAlaPheIysGlySer 777
Db 2567 CACTGCCATGTTGGCGTCAATTTACTGCATACCTAAACGGCGCCCTTCAAAGTTCC 2626
Qy 778 AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVal 797
Db 2627 GCACCTCTCAATCTTTCGATCGTCTTCCAGGTTCTTCGGTGGGCTCTTGGTTCAAT 2686
Qy 798 GluIlePhePheSerArgHisSerProLeuLeuTrpGlyTrpIysAsnGlyAsnLeuIys 817
Db 2687 GAAATTTTCTTCAGCAACCACTGCTCTCTGGTATGGGTAT---GGTGGTGGAGTAAG 2743
Qy 818 TrpLeuGluArgPheAlaTrpIleAsnThrThrIleTrpPheThrSerLeuProLeu 837
Db 2744 TTCCTGGAAGTTTTCGTACATTTAACTCCATCGTATACCTTGGACATCTATCCCGTCT 2803
Qy 838 LeuAlaTrpCysThrLeuProAlaValCysLeuLeuThrGlyIysPheIleMetProSer 857
Db 2804 TTGGCTATTGCATTTGCTGCCATCTGCTTGTGTGACAGGAAATTTTATCACGCCAGAG 2863
Qy 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2864 CTTAAACATGTTGCCAGCTCTGTTGTCATGTCACCTTTTCTCATCTGCACTTTTGTACGAGC 2923
Qy 878 IleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPhe 897
Db 2924 ATCTGGAATGAGATGGAGTGGTGTAGGCATGACTGCTGTGGGAAGCAAGACGATTT 2983

QY 216 ArgMetAspAspTyrLysSerLysGln-----GlyLeuGlyGly 229
DB 801 CGGATGGAGAAATGGAAGCAGACAGACAGAGAGATGCACACAGCGGGAATGATGGTGT 860
QY 230 GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnPro 249
DB 861 GGTGATGGTGCAC-----GATGCTGATCTACCACTAATGGATGAACAGACAGACNA 914
QY 250 LeuSerArgLysValSerLysLeuAlaSerSerLysValAsnProTyrArgMetValLeuVal 269
DB 915 CTGTCAGGAAAATCCCACTTCATCAAGCCAGATTAATCCATATAGGATGATTAATCAT 974
QY 270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgLysLeuHisProValProAsp 289
DB 975 ATTCGGCTTGTGGTGTGGGTCTCTTCCACTACCGAGTGCATCCGGTGAATGAT 1034
QY 290 AlaIleGlyLeuTyrLeuValSerLysLeuGlyLeuTyrPheAlaIleSerTyrPhe 309
DB 1035 GCATTTGCTTGTGGCTCATCTGTATCTGTGAAATCTGGTTGCCATGCTTGGATT 1094
QY 310 LeuAspGlnPheProLysTyrPheProLysArgGluThrTyrLeuAspArgLeuSer 329
DB 1095 CTTGATCAATTCCTCAAGTGGTTCCTATTGAGAGAGACATTAAGAGCGGCTGTCA 1154
QY 330 LeuArgTyrGluArgGluGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
DB 1155 CTGAGGTTGACAGCAAGAGCGGAGCCATCTCAACTTGTCTCAAAATGATTTCTTTCAGT 1214
QY 350 ThrValAspProLeuLysGluProLeuValThrAlaAsnThrValLeuSerLysLeu 369
DB 1215 ACGGTTGATCCCTTAAGGAACCTCTTGTGGTCAACAACTACTGTTCTATCTATCTT 1274
QY 370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 389
DB 1275 TCGGTGGATTAATCTCTGATTAAGGTCTCTCTGATGTTCTGATGATGGTCTCAATG 1334
QY 390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTyrValProPheCys 409
DB 1335 CTAAAGTTGAAGCAATTAATCTGAACATCTGAATTTGCAAGAAATGGTCTCTTCTGC 1394
QY 410 LysLysPheGlyLeuGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
DB 1395 AAACGGTACNATTTGAACCTCGGCTCCAGAGTGTACTTCCACAGAGATAGACTAC 1454
QY 430 LeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyr 449
DB 1455 TTGAAGACAGCAAGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGAT 1514
QY 450 GluGluPheLysValArgLysAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
DB 1515 GAGGAATTCAGGTGAGAAATCAATGCCTTAGTTGCCAAAGCCAGAAAGTTCTCTGAAGAA 1574
QY 470 GlyTyrPheLeuMetLysAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 489
DB 1575 GGATGGACAAATGCAAGATGGAACCCCTGGCTCGGAAACAATGTTCTGGATCATCTCTGGA 1634
QY 490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
DB 1635 ATGATTCAGGTCTCTCTGGCCAAAGCGAGGSCCTTGACTGTGAGGGAATGAACCTGCCA 1694
QY 510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 529
DB 1695 CGATTTGGTTTATGTTTCTAGAGAGAAACGACCGAGCTATAACCATCAATAAGAAAGCTGGT 1754
QY 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
DB 1755 GCTATGANTGANTGGTCCGAGTCTCTGCTGACTTAACAATGCTCCATATTTGTTAATC 1814
QY 550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
DB 1815 TTGGATTGTGATCACTACATCAACACAGCAAGGCTATAAGAGAGCAATGTTTATG 1874
QY 570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589

DB 1875 ATGGACCCCTTACTAGGAAGAAGGTTTGTCTATGTACAGTTCCTCAAGATTTGTATGGG 1934
QY 590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspTyrIleAsnMetLys 609
DB 1935 ATTGATCCCATGACCGATATGCTAACCGGAATGTGTCTTTTTTGATATCAACATGAAA 1994
QY 610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln 629
DB 1995 GGTTTGGATGTTATTCAGGGTCCAAATTTATGTTGTGTACTGGATGTGTTATTAAGAGGAG 2054
QY 630 AlaIleuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThrCysAsp 649
DB 2055 GCATTTATGTTATGATGATGCCCCCAAA--ACAAAGAGCCACCATCAAGGACTCTGCAAC 2111
QY 650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
DB 2112 TGCTGGCCCAAGTGTGTCTTTTGTCTGTCTTGTGCAATAGGAAGCAAAAGAACT 2171
QY 663 LysAspGlyLeuProGlu----- 668
DB 2172 ACCAAACCCAAACAGAGAGAAAGTTATTTTCAAGAAAGAGAGAACCAATCC 2231
QY 669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
DB 2232 CTGCATATGCTCTTGGTGAATTTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAGGCC 2291
QY 681 MetLeuSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
DB 2292 GGTATGTAAATCAACAAAAATTAGAAAGAAATTTGGCCAATCTTCTGTTTTGTGTACA 2351
QY 701 SerThrLeuMetGluGluGlyValProProSerSerProAlaAlaLeuLeuLys 720
DB 2352 TCCACACTTCTCGAGAAATGGTGAACCTTGAAGAGTGAAGTCTCTCTCTTTTGA 2411
QY 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTyrGlyLeuGluLeu 740
DB 2412 GAAGCTATACATGATCATTTAGTTGTGGTTATGAAGACAGACAGACTGGGGAAAAGAGAT 2471
QY 741 GlyTyrPheLeuTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
DB 2472 GGCTGGATCATGATGATCATGATTTACAGAGATATTTCTAAGTGTTCAGATGATGATG 2531
QY 761 GlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
DB 2532 GGTGGGGGTCAAATTTACTGATACCTAAACGGGTGTCATTAAGAGTCTTGACACTCTG 2591
QY 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIlePhe 800
DB 2592 AATCTTTCAGATCGTCTTCCAGGCTGCTCGGTGGGTCTTGGGTCTTATTGAGATCTTC 2651
QY 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTyrLeuGlu 820
DB 2652 TTCAGCAATCATGCCCTCTTTGGTATGGGTAT---GGTGGGGTCTCGAAATTTTGGAA 2708
QY 821 ArgPheAlaTyrIleAsnThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
DB 2709 AGATTTCTCATCACTCACTCCATCGTGTATCTTGGACATCTATTCCCTCTTGCTTAC 2768
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
DB 2769 TGTCATTTGCCCTGCATCTGTTTATTGACAGGGAATTTATCACTCCAGAGCTGAATAAT 2828
QY 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
DB 2829 GTTCCAGCCCTGGTTCATGTCATCTTTTATCTGCAATTTTGTACAGGATCTCTAGAA 2888
QY 881 MetArgTyrSerGlyValSerIleGluGluTyrTyrArgAsnGluGlnPheTyrValIle 900
DB 2889 ATGAGATGGAGTGGTGTGGAAATGATGACACTGGTGGAGGAATGAGCACTTCTGGGTCA 2948
QY 901 GlyGlyValSerAlaHisLeuPheAlaValGlnGlyLeuLeuLysValLeuAlaGly 920

Db 2949 GGAGGTGTCTCCACACCTCTTTGCTGTGTTCCAGGGACTTCTCAAGGTCTATAGCTGGT 3008
 Qy 921 IleAspThrAsnPheThrValThrSerIysAlaThrGlyAspGluAspAspGluPheAla 940
 Db 3009 GTTGATCAAGCTTCCCGTGACATCAAG-----GGTGGAGATGATGAGGAGTTCTCA 3062
 Qy 941 GluLeuTyRAlaPheIysTrpThrThrLeuLeuIleProProThrThrThrLeuLeuIle 960
 Db 3063 GAGCTATATACATTCAATGAGACTACCTTATTGATATCTCTACACCTTGCTTCTATTG 3122
 Qy 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyIleGlnSerTrp 980
 Db 3123 AACTTCATTGGTGGTGGTGGTGGTTCAAATCGCATCAATACGATATAGTATGATGG 3182
 Qy 981 GlyProLeuPheGlyIysLeuPhePheAlaPheTrpValIleValIleLeuTyRProphe 1000
 Db 3183 GGCCCCCTCTTTGGGAAGCTATTCTTTTGCAATTTTGGGTGATTTGCCATCTTTATCCCTTT 3242
 Qy 1001 LeuIysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
 Db 3243 CTCAAAGTTTGGTTGAAGCAAAACAGACACCAACGATGTTCATCGTCTGCTGCATTT 3302
 Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPropheIleValArgThr 1040
 Db 3303 CTGCTGGCTTCAATCTCTCGCTCTCTTTGGGTTCGATTCGATCTTCTTCTTGGCAGGAT 3362
 Qy 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db 3363 GATGTCGCTCTCTTGGAGAGTGTGGTTGGATTGC 3398
 RESULT 6
 AAZ99518
 ID AAZ99518 standard; DNA; 3969 BP.
 AC AAZ99518;
 DT 03-JUL-2000 (first entry)
 DE DNA encoding a maize cellulose synthase.
 KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 OS transgenic plant; plant breeding marker; ss.
 FH Zea mays.
 FT Key
 FT CDS
 FT Location/Qualifiers
 FT 144..3398
 FT /*tag= a
 FT /product= "cellulose synthase"
 FT /note= "no termination codon given"
 FN WO200009706-A2.
 XX 24-FEB-2000.
 XX 16-AUG-1999; 99WO-US018760.
 XX 17-AUG-1998; 98US-0096822P.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 XX WPI; 2000-224343/19.
 XX P-PSDB; AAY84116.
 XX New genes which encode maize cellulose synthase polypeptides in plants
 XX useful for modulating the expression of cellulose synthase in plants and
 XX PT to produce transgenic plants expressing the novel protein.
 XX Claim 1; Page 152-157; 119pp; English.
 PS The present sequence encodes a maize cellulose synthase polypeptide. The
 CC

CC cellulose synthase can be used for the improvement of stalk quality for
 CC improved stand or silage. It also provides an increased concentration of
 CC cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting or
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein
 XX
 SQ Sequence 3969 BP; 1022 A; 838 C; 1018 G; 1091 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3969
 Score: 4073.50 Matches: 757
 Percent Similarity: 79.68% Conservative: 129
 Best Local Similarity: 68.08% Mismatches: 139
 Query Match: 72.12% Indels: 87
 DB: 3 Gaps: 15
 US-10-627-132-30 (1-1052) x AAZ99518 (1-3969)
 Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgGlnLeuValLeu 20
 Db 144 ATGAGGCGAGCGCGCGCTGGTGGCGGCTCCACAAACCGACAGCTCGTGTGATC 203
 Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
 Db 204 CGCGCGACGGCGATCCCGGGCGAGCCCGGGGAGCAGAACCGGAGGTGTGCCAG 263
 Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
 Db 264 ATTTGGCGCAGCAGCTGCGCTTGGCCCGGGGAGCCCTTCGTGGGTGTCACAGAG 323
 Qy 59 CysGlyPheProValCysArgProCysTyRLeuArgGlyGluArgGlyThrGluAsn 78
 Db 324 TGGCGCTTCCCGCTGTCGGCGGACTGCTAGATACAGAGCGCTCAAGGGCTGCCAACCGTGACCGGTGAC 383
 Qy 79 CysProGlnCysIysThrArgTyRLeuArgLeuIysGlySerProArgValAlaGlyAsp 98
 Db 384 TGCCCCCAGTGCAGACTCGATACAGCGCTCAAGGGCTGCCAACCGTGACCGGTGAC 443
 Qy 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
 Db 444 GAGGAGGAGGAGCGCGTGGATGACCTGGCAACAGATTCAACTGGGACGGGCATGACTCG 503
 Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
 Db 504 CAG-----TCTGTGGCGGAGTCCATCTCTACGGC 533
 Qy 139 ArgMetSerTyRGlYArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
 Db 534 CACATGAGTACGCGCGCTGGA-----GGTGACCTTAATGGCGGCCCAAGCTTTC 584
 Qy 159 ProIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
 Db 585 CAGCTC-----AACCCCAATGTTCCATC-----CTCACCACCGG 620
 Qy 179 TyrGlyHisGlyGluValSerSerSerLeuHis----- 189
 Db 621 CAATGTGGATGACATCCCAACCGGAGCAGCACCGGTGGTCTTTTTCATGGTGGT 680

QY 190 -----LysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db 681 GGGGAAAGAGGATACATCCCTTCTTATGCGGATCCCGATACCTGTGTGCAACCCAGG 740
QY 204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Db 741 TCTATGAGCCATCCAGGATCTTCTGTCATATGGTATGGTATGGTGTGGAGGAA 800
QY 216 ArgMetAspTrpLysSerLysGln-----GlyIleLeuGlyGly 229
Db 801 CGGATGGAGAAATGGAAGCAGACAGACAGAGAGGATGCCACGCGGGAATGATGGTGT 860
QY 230 GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnPro 249
Db 861 GGTATGATGGTGAC-----GATGCTGATCACCACCTAATGGATGAAGCAGACACAA 914
QY 250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
Db 915 CTGTCCAGGAAATCCATCCATCAAGCCAGATTATCCATATAGATGATATCATTT 974
QY 270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
Db 975 ATTCGGCTTGTGGTGTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGAT 1034
QY 290 AlaIleGlyLeuTrpLeuValSerIleLeuCysGluIleTrpPheAlaIleSerTrpIle 309
Db 1035 GCATTTGCTGTGCTCATCTCTATCTGTGAAATCTGTGTTTGGCCATGTCTTGGATT 1094
QY 310 LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTrpLeuAspArgLeuSer 329
Db 1095 CTGTGATCAATCCCAAGTGTCTTCTATTTGAGAGAGACTTACTAGACCGGTGTCA 1154
QY 330 LeuArgTyrGluArgGluGlyProSerLeuLeuSerAlaValAspLeuPheValSer 349
Db 1155 CTGAGGTTTCGACAGAGAGCCAGCCATCTCAACTTGTCTCAATTTGTTCTTGTCTAGT 1214
QY 350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu 369
Db 1215 ACGGTGTGATCCCTTAAAGAAACCTCTTGTGTGACAAACAAATCTGTTCTATCTCTT 1274
QY 370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGluValSerMet 389
Db 1275 TCGGTGGATATCTCTGTGATNAGGTTTCTGTCTATGTTCTGATGATGGTGTCTGCAATG 1334
QY 390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys 409
Db 1335 CTAACGTTTGAAGCATTTATGAAACATCTGAATTTGCAAGAAATGGTTCCTTTCTGTC 1394
QY 410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
Db 1395 AAACGGTCAATATTTGAACCTCGCGCTCCAGAGTGTGTCTTCCACAGAGATAGACTAC 1454
QY 430 LeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyr 449
Db 1455 TTGAAGAACAGGTGGCAGAACCTTTGTTAGGAGAGAGAGCAATGAAGAGAGATAT 1514
QY 450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
Db 1515 GAGGAATTCAGGTGAGATCAATGCCITAGTTGTCMAAGCCAGAAAGTTCCTGAAGAA 1574
QY 470 GlyTrpIleMetLysAspGlyThrProTrpProGlyLysAsnThrArgAspHisProGly 489
Db 1575 GGATGGCAATGCAAGATGGAACCCCTCGCGCTCCAGAGTGTGTCTTCCATCATCTCTGGA 1634
QY 490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
Db 1635 ATGATTCAGGTCTCTCTTGGCCAAAGCGGAGCGCTTGACTGTGAGGAAATGAATGCTCA 1694
QY 510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 529
Db 1695 CGATTTGGTTTATGTTCTTAGAGAGAAACGACCGGCTATACCATCATCAAGAAAGCTGT 1754
QY 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549

Db 1755 GCTATGAATGATGGTCCGAGTCTCTGCTGTACTAACAAATGTCTCCATATTTGTTAAAC 1814
QY 550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
Db 1815 TTGGATTGTGATCACTACATCAACACAGCAGGCTATTAAGGAAGCAATGTGTTTATG 1874
QY 570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
Db 1875 ATGGACCTTTACTAGGAAGAAGGTTTGTCTATGTACAGTTCCCTCAAGATTTGATGGG 1934
QY 590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
Db 1935 ATTGATCCCGATGACCGGATGCTAACCGGAATGTGTCTTTTGTATCAACATGAAA 1994
QY 610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln 629
Db 1995 GGTTTGGATGTTATTCAGGTCCTAATTTATGTGTGCTGATGTGTTATTAAGGCGCAG 2054
QY 630 AlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThrCysAsp 649
Db 2055 GCATTTATGTTTATGATGCCCCAAA---ACAAAGAGCCACCATCAAGACTTGTCAAC 2111
QY 650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
Db 2112 TGCTGGCCAAAGTGGTGTCTTGTGCTGTGCTTGGCAATAGGAAGCAAAAGAACT 2171
QY 663 LysAspGlyLeuProGlu----- 668
Db 2172 ACCAAACCCAAAACAGAGAGAAAAGTTATTTTCAAGAAAGAGAGAACCAATCC 2231
QY 669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
Db 2232 CCGTCATATGCTCTTGGTGAATTTGACGAAGCTGCTCCAGCAGCTGAGAAAGAGGCC 2291
QY 681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
Db 2292 GGTATTTGAAATCAACAAAAATTAGAAAGAAATTTGGCCAAATCTCTGTTTGTGTACA 2351
QY 701 SerThrLeuMetGluGluGlyValProProSerSerProAlaAlaLeuLeuLys 720
Db 2352 TCCACACTTCTCGAGAAATGGTGGAACTTGAAGAGTGCAGTCTCTGCTCTTTTGAAA 2411
QY 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
Db 2412 GAAGCTATACATGTCATTTAGTTGTGTTATGAAGACAGACAGACTGGGGAAGAGATT 2471
QY 741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
Db 2472 GGCTGGATCTATGATCAGTTACAGAGATATTCTAACTGGTGTTCAGATGCATTTGCAT 2531
QY 761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
Db 2532 GGTGGCGGTCAATTTACTGCATACCTAAACGGGTGCAATTCAAAGGTTCTGCACCTCTG 2591
QY 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
Db 2592 AATCTTTTCAGATCGTCTTCCAGCGTCTCGTGGGCTCTTGGTCTCTATTGAGATCTTC 2651
QY 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
Db 2652 TTCAGCAATCATTCCTCTTGTGTATGGGTAT---GGTGGCGGTCTCGAAATTTTGGAA 2708
QY 821 ArgPheAlaTyrIleAsnThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
Db 2709 AGATTTTCTACATCAATCCATCGTATCTCTGGACATCTATCTCCCTCTTGGCTTAC 2768
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTACATTCCTGCCATCTGTTTATTGACAGGAAATTTTATCACTCCAGAGCTGAATAAT 2828
QY 861 PheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880

```
Db 2829 GTTCCAGCCTGTGGTTTCATGTCACCTTTTATCTGCATCTTTTGTCTACGAGCATCTAGAA 2888
Qy 881 MetArgTrpSerGlyValSerIleGluGluTrpTrpArgHsnGluGlnPheTrpValIle 900
Db 2889 ATGAGATGGAGTGGTGTGGAAATGATGACTGGTGGAGGAATGAGCAGTCTCTGGGTCAAT 2948
Qy 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlnGlyLeuLeuLysValLeuAlaGly 920
Db 2949 GGAGGTGTCTCCTCACACCTCTTTGCTGTGTTCAGGGACTTCTCAAGGTCAATAGTGTGT 3008
Qy 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
Db 3009 GTTCATCAAGCTTTCACCGTCACATCAAG-----GGTGGAGATGATGAGGAGTCTCTCA 3062
Qy 941 GluLeuTrpAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIle 960
Db 3063 GAGCTATATACATTCAAATGGAGTACCTTATGTATACCTCTCTACACCTTGTCTTCTATTG 3122
Qy 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTrpGlnSerTrp 980
Db 3123 AACTTCATTGGTGTGGTCTGGCTTCAATCGGATCAATAACGGATATGATGATGATG 3182
Qy 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTrpProphe 1000
Db 3183 GGCCCTCTTTGGGAAGCTATTCTTTGCAATTTGGGTGATTGTCATCTTATCCCTTT 3242
Qy 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
Db 3243 CTCAAAGGTTTGGTGGAGGCAAAACAGGACACCAACGATTTGTCATCGTCTGGTCCATT 3302
Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 3303 CTGCTGGCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1060
Qy 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3363 GATGTCGGCTCTCTGAGGAGTGTGGTTTGGATTGC 3398

RESULT 7
AAZ99503
ID AAZ99503 standard; DNA; 3969 BP.
XX
AC AAZ99503;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
CDS 144..3398
FT /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
XX
FN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI; 2000-224343/19.
DR P-PSDB; AAY84111.
```

```
XX New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 113-118; 119pp; English.
XX
```

```
XX The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
```

SQ Sequence 3969 BP; 1022 A; 838 C; 1018 G; 1091 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3969
Score:	4073.50	Matches:	757
Percent Similarity:	79.68%	Conservative:	129
Best Local Similarity:	68.08%	Mismatches:	139
Query Match:	72.12%	Indels:	87
DB:	3	Gaps:	15

US-10-627-132-30 (1-1052) x AAZ99503 (1-3969)

Qy	1	MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgHsnGluLeuValLeulle	20
Db	144	ATGAGGCGAGCGCGGGCTGGTGGCGGCTCCACAAACGCGACGAGCTGTCGTGCATC	203
Qy	21	Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu	38
Db	204	CGCGCGACGCGGATCCCGGCGGAGCCCGCGGGAGCAGAACGGCGAGGTGTGCAG	263
Qy	39	IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu	58
Db	264	ATTTCGCGCAGCAGCTGCGCTTGCCTTCCCGGGGGACCCCTTCGTGGCGGTGCAACGAG	323
Qy	59	CysGlyPheProValCysArgProCysTrpGluArgArgGluGlyThrGlnAsn	78
Db	324	TGGCCCTTCCCGTCTCGCGGACTGCTACGAATACGAGCGCGGGAGGCGACGAGAAC	383
Qy	79	CysProGlnCysLysThrArgTrpLysArgLeuLysGlySerProArgValAlaGlyAsp	98
Db	384	TGCCCCCAGTGCAAGACTCGATACAAAGCGCTCAAGGGCTGCCCAACGTGTGACCGGTGAC	443
Qy	99	AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln	118
Db	444	GAGGAGGAGGACGGCGTCGATGACCTGGACACAGATTCACTGGGACGGCCATGACTCG	503
Qy	119	GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly	138
Db	504	CAG-----TCTGTGCCGAGTCCATCTCTACGCG	533
Qy	139	ArgMetSerTrpGlyArgGlyProAspAspGlyAsnAsnThrProGlnIlePro	158
Db	534	CACATGAGCTACGCGCGTGA-----GGTGACCTTAATGGCGGCCCAAGCTTTC	584


```
Db 2709 AGATTTTCTACATCAACTCCATCGTGTATCTCTGGACATCTATTCCCTCTTGGCTTAC 2768
Qy 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTACATTTGCTGCGCATCTGTTATTGACAGGGAATTTATCACTCCAGAGCTGAATAAT 2828
Qy 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 2829 GTTGCCAGCCTGTGGTTCATGTCACCTTTTATCTGCAATTTTGTCTAGCAGATCCTAGAA 2888
Qy 881 MetArgTrpSerGlyValSerIleGluGluTrpTrpArgAenGluInPheTrpValIle 900
Db 2889 ATGAGATGGAGTGTGTGTGAATTTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCAAT 2948
Qy 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
Db 2949 GGAGGTGTCTCCTCACACCTCTTTGCTGTGTTCAGGGACTTCTCAAGGTTCATAGCTGGT 3008
Qy 921 IleAspThrAenPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
Db 3009 GTTGATCAAGCTTCCCGTGACATCAAG-----GGTGGAGATGATGAGGATTTCTCA 3062
Qy 941 GluLeuTyraIaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIle 960
Db 3063 GAGCTATATACATTTCAATGGACTACCTTATTGATACCTCTACACCTTGCTTCTATTG 3122
Qy 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAenAenGlyTyraGlnSerTrp 980
Db 3123 AACTTTCATTGGTGTGGTGGCTGGCTTTCAAATCCGATCAATAACGATATGAGTCATGG 3182
Qy 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyraProPhe 1000
Db 3183 GGCCCCCTCTTTGGGAAGCTATCTTTGCAATTTGGGTGATGTGTCATCTTATCCCTTT 3242
Qy 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrpSerIle 1020
Db 3243 CTCAAGAGTTTGGTTGGAAGGCAAAACAGGACACCAACGATTTGTCATGCTGTGTCATT 3302
Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 3303 CTGCTGGCTTCAATCTTCTCGCTCTCTTTGGGTTCGGATTGATTCCTTTCTTGGCAAGAT 3362
Qy 1041 LysGlyProAspValArgGlnCysGlyIleAenCys 1052
Db 3363 GATGTCCTGCTTCTTGAGAGTGTGGTTTGATGTC 3398

RESULT 8
ID AA299491 standard; DNA; 3568 BP.
AC AA299491;
XX
DT 03-JUL-2000 (first entry)
DE DNA encoding a maize cellulose synthase.
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 63..3236
FT /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
```

```
PR 17-AUG-1998; 98US-0096822P.
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Dhuga KS, Helentjaris TG, Bowen BA, Wang X;
PI P-PSDB; AAY84107.
XX WPI; 2000-224343/19.
DR
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
XX Claim 1; Page 82-87; 119pp; English.
XX
XX The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 3568 BP; 879 A; 825 C; 941 G; 922 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 0 Length: 3568
Score: 4005.00 Matches: 743
Percent Similarity: 79.04% Conservative: 128
Best Local Similarity: 67.42% Mismatches: 141
Query Match: 70.91% Indels: 90
DB: 3 Gaps: 16

US-10-627-132-30 (1-1052) x AA299491 (1-3568)
Qy 14 ArgAsnGluLeuValLeuIleArg-----GlyHisGluAspProLysProLeuArgAla 31
Db 12 CGTCCGAGTCTGTCGTCTCCGCGATGGCGAGCCAGGCCGAGCCCATGGACCAG 71
Qy 32 LeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAsp 51
Db 72 CGGAACGCGCAGGTGTGCCAGATTGGCGGCACGACGTCGGCGCAACCCCGCGGGGAG 131
Qy 52 LeuPheValAlaCysAenGluCysGlyPheProValCysArgProCysTrpGluTrpGlu 71
Db 132 CCTTTCGTGGCTTCAACGAGTGGCCCTTCCCCATCTGCCGGGACTGCTAGTACGACG 191
Qy 72 ArgArgGluGlyThrGlnAenCysProGlnCysLysThrArgTrpLysArgLeuLysGly 91
Db 192 CGCCGCGAGGCGCAGCAACTGCCCCAGTGCAGACCCGCTTCAAGCGCTTCAAGGGG 251
Qy 92 SerProArgValAlaGlyAspAspGluGluAspIleAspAspLeuHisGluPhe 111
Db 252 TCGCGCGCGTGTCCCGGGACGAGAGGAGCGCGCTCGACGACCTTGGAGAACGAGTTC 311
Qy 112 AsnIleAspAspGluAenGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIle 131
Db 312 AACTGGAGCGCAAGACGACGACTCCAGTACCTC----- 344
```


QY 132 ThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAsp---AspGlyAsp 150
Db 345 GCCAGTCTCATGCTCCACCCACATGAGCTACGGCCGGCCGCGCATCGACGCGGTG 404
QY 151 GlyAsnAsnThrProGlnIleProIleThrGlySerArgSerValProValSer 170
Db 405 CGCAGCCATCCACCCCATCCC-----ATGTTCCCTC--- 440
QY 171 GlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerLeuHis--- 189
Db 441 -----CTCACAACCGACAGATGGTCGATGACATCCCGCGGACCGACGCGC 488
QY 190 -----lysArgIleHisProTyrProValSerGlu 199
Db 489 CTTGTGCCCTCGTTCGTGGTGGCGGGGAGAGGATTCACCCCTCCCGTACCGGAT 548
QY 200 ProGly-----SerAlaLysTrpAspGluLysGlu----- 210
Db 549 CCCAACCTTCGTGCAACCGAGGTCATGGACCCCTCCNAGGATCTCGCCGCATATGCC 608
QY 211 -----ValSerTrpLysGluArgMetAspTrpLysSerLysGlnGlyIleLeu 227
Db 609 TACGGGAGCGTAGCATGAAGAGAGGATGGAGAGCTGGAAGCAGAGCAGGAGGATG 668
QY 228 -----GlyGlyGlyAlaAspProGluAspMetAspAlaAspValAla 241
Db 669 CACCAGACGAGGAACGATGGCGCGCGCGATGGTGAT-----GATGCAGATCATCCA 722
QY 242 LeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLysVal 261
Db 723 CTAATGGATGAAGCTAGACAGCATGTGTCAGAAAGATCCCGCTTCCTTCAAGCCAAATC 782
QY 262 AsnProTyrArgMetValIleValArgLeuValValLeuAlaPhePheLeuArgTyr 281
Db 783 AACCCCTATAGGATGATTATAATAATTCGGCTAGTGGTTTGTGTTCTTCTTCCACTAC 842
QY 282 ArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGlu 301
Db 843 CGAGTGATGATCCCGGCGCGTATGATGCTTTTATGGCTCATATCTGTGATCTGTGAA 902
QY 302 IleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAspArg 321
Db 903 ATTGGTTGGCATGCTCTTGGATCTTGACCGATTTCCAAAGTGGTTTCTTATCGAGAG 962
QY 322 GluThrTrpLeuAspArgLeuSerLeuArgTyrGluArgGlyGluProSerLeuLeu 341
Db 963 GAAACCTATCTTGACCGGCTGAGTTTAAGTTTGACAAGGAGGSCATCTTCTCAACTC 1022
QY 342 SerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuValThr 361
Db 1023 GCCCTGTGATTTCTTTGTCACTACGGTGTGATCCCTTTGAAGGAACCTCCATTTGTCAC 1082
QY 362 AlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyr 381
Db 1083 GCTAATACGTCTATCATCTTTCGGTGGATATATCCAGTTGATAAGGTTTCATGCTAC 1142
QY 382 ValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPhe 401
Db 1143 GTTCTGATGATGGTGGCTGCCATGCTGACATTTGAAGCATGCTCTGAACATCTGAATTT 1202
QY 402 AlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAlaProGluPhe 421
Db 1203 GCAAAGAAATGGGTTCTCTTCTGCAAAAAGATATAGCTTTGAGCCCTCGTGTCCAGAGTGG 1262
QY 422 TyrPheSerLeuLysValAspTyrIleLysAspLysValGlnProThrPheValGlnGlu 441
Db 1263 TACTTCCAAACAGAGATAGCTTACCTGAAGACAAGGTTGGCGCCAAACTTTGTTAGAGAA 1322
QY 442 ArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuValAla 461
Db 1323 CGGAGAGCAATGACAGAGAGATGATGAGGAATTCAGAGTTCAGATCAATCAATGCTTGTGCT 1382
QY 462 LysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProGly 481

Db 1383 AAAGCCCAAAAGTTCTTGAGGAAGATGGACAATGCGAGGATGGAACTCCATGGGCCGGA 1442
QY 482 AsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHis 501
Db 1443 AATAATGTCGCGTATCATCTCTGGATGATTCAGGTTTCTTGGTCAAGAGTGGTGGCCAT 1502
QY 502 AspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGly 521
Db 1503 GATGGGAAGAAATGAGCTGCTCGATTGTTTATGTTTCAAGAGAAAAACGGCAGGC 1562
QY 522 PheGlnHisLysLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeu 541
Db 1563 TACAACCATCAAGAAGGCTGGTCTATGAATGCAATGGTGGTCCGAGTCTCTGCTACTA 1622
QY 542 ThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAla 561
Db 1623 ACTAATGCTCTTATTTGCTGACTGATGATGATGATGATGATGATGATGATGATGATG 1682
QY 562 IleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyrVal 581
Db 1683 ATAAAGGAAGCAATGTTTATGATGATGATGATGATGATGATGATGATGATGATG 1742
QY 582 GlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThr 601
Db 1743 CAGTTTCTCAAGATTTGATGGGATGATCGCCATGATGATGATGATGATGATGATGAT 1802
QY 602 ValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGly 621
Db 1803 GTCTTTTTCGATATCAACATGAAAGTTGGATGGTATCCAGGGCCCAATTTATGTTGG 1862
QY 622 ThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLys 641
Db 1863 ACTGATGTGCTTTCAGAAAGCAGGCATATATGCTACGATGCTCCCAAA---ACAAG 1919
QY 642 ArgProLysMetValThrCysAsp-----CysCysProCysPhe 654
Db 1920 AAGCCACCATCAAGAATCTGCACTGCTGGCCAAAGTGGTGCATTTGCTGTGCTGTTT 1979
QY 655 Gly---ArgLysLysArgLysHisAlaLysAspGlyLeuProGlu----- 668
Db 1980 GGTAAACAGAGAGACCAAGAGAACCCAGACCTCTAAACCTTAATTTGAGAGATAAG 2039
QY 669 -----GlyThr 670
Db 2040 AAACCTTTTAAAGAAAAAGAAAAATCAAGCCCTCGCATATGCTCTTGGTGAATTCATGAA 2099
QY 671 AlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLys 690
Db 2100 GCCGCTCCAGAGCTGAAAATGAAAAGGCTAGTATTTGTAATCAACAGAAAGTTGGAAG 2159
QY 691 ArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValPro 710
Db 2160 AAATTTGGCCAGTCTTTCAGTTTGTTCATCCACACTTCTTGAGATGGTGGAAACCTG 2219
QY 711 ProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 730
Db 2220 AAGAGTGCCAGTCTCTCTCTCTGAAGGAAGCATATACATGTCATCAGTTGCGATAT 2279
QY 731 GluAspLysThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAsp 750
Db 2280 GAAGACAAAACAGCTCGGGGAAAAGATATTGTTGGATTTATGGATCATGTCACAGAAGAT 2339
QY 751 IleLeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLys 770
Db 2340 ATTCTTACTGGGTTTAAAGTGCATGCTCCATGCTGGTGGCGGCTCAATTTACTGCATCTAA 2399
QY 771 ArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 790
Db 2400 CGGGCCGCTTCAAAGGTTCCGACCTCTCAATCTTTCCGATCGTTTTCACCAGGTTCTT 2459
QY 791 ArgTrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGly 810

Db 569 GACGGCGCCGCCACAGAGAC-----GACCCGAGTACGTCGCCGAGTCC 613
Qy 135 MetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGlyAspGly---AsnAsn 153
Db 614 ATGCTCAGGCGCGCAGATAGCTACGGCGCGCGCGCGCGCCAGCCCGCTTCAGCC 673
Qy 154 ThrProGlnIleProProIleIleThrGlySerArgSerValProValSerGlyGlu--- 172
Db 674 GTCCCAACAGTCGCGCTCCTCCACCAACGCGCCAGATGGTTGATGATGATCCCGCGGAGCAG 733
Qy 173 -----PhePro-----IleThrAsnGlyTyrGlyHisGlyGluValSerSer 187
Db 734 CACGCGCTCGTCCCTCTACATGAGCGCGCGCGCGCGCGCGCG----- 778
Qy 188 LeuHisLysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db 779 -----AAGAGGATCCACCGCTCCTTCGCGAGATCCCAACCTTCCAGTGCAACCGAGA 832
Qy 204 LysTrpAspGluLysGlyGlu-----valSerTrpLysGlu 215
Db 833 TCCATGGAGCCCGTCCCAAGGATCTGCCCGCTACCGGATATGGCAGCGTGGCTGGAGGAG 892
Qy 216 ArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGly 230
Db 893 AGATGGAGGCTGGAAGCAGAGCAGAGCGCGCTGCAGCATGTTCAGGAGCGAGGCTGGC 952
Qy 231 AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaAargGlnProLeu 250
Db 953 GGTGATTGGGATGGCGACGATCGATCTGCCACTAATGGATGAAGTAGCGAGCCATTG 1012
Qy 251 SerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleValVal 270
Db 1013 TCCAGAAAGTCCCTATATCAAGCCGAATTAATCCCTACAGGATGATTATCGTTATC 1072
Qy 271 ArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisAspProValProAspAla 290
Db 1073 CGTTTGGTGGTTTGGGTTCTTCCATCCAGGATGATGCATCCCGCGGAAAGATGCA 1132
Qy 291 IleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeu 310
Db 1133 TTTGATGTGTGCTCATATCTGTAATCTGTGAATCTGTTTGGATGTCTCGATTCCT 1192
Qy 311 AspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLysSerLeu 330
Db 1193 GATCAGTTCCTCCAAAGTGGCTTCCATCGAGAGAGAGACTTACCTGGACCGTTTGTCTA 1252
Qy 331 ArgTyrGluArgGluGluProSerLeuLeuSerAlaValAspLeuPheValSerThr 350
Db 1253 AGTTTGAAGGAAGGTCAACCCCTCTCAGCTTGTCTCCATCGACTTCTTTGTGAGTACG 1312
Qy 351 ValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAla 370
Db 1313 GTTGATCCCAAGGAACCTCCCTTGTGTACAGCGAAGACACTGTCTTCCATCTCTTCT 1372
Qy 371 ValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeu 390
Db 1373 GTGATTAATCCGGTTGAGAAGTCTCTCTGCTATGTTTCTGATGATGGTGTCTCAATGCTT 1432
Qy 391 ThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTyrValProPheCysLys 410
Db 1433 ACCTTTGAAGCATGTCTGAAACATCTCAATTTGCAAGAATGGTTTCTTTCAGCAAA 1492
Qy 411 LysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeu 430
Db 1493 AAGTTTAATATCGAGCTCGTCTCTGAGTGGTACTTCCACAGAGATAGACTACCTG 1552
Qy 431 LysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGlu 450
Db 1553 AAAGACAGGTTGCTGCTTCAATTTGTTAGGAGAGAGGCGCGATCAAGAGAGATACGAG 1612
Qy 451 GluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGly 470

Db 1613 GAATTCAGGTAAGATCAATGCCCTTGGTGTCAAAAGCCCAAAAGGTTCTCTGAGGAAGA 1672
Qy 471 TrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMet 490
Db 1673 TGGCAATATGCAAGATGGAAGCCCTTGGCTTGGAAACAACGTACGCGATCATCTCTGAATG 1732
Qy 491 IleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArg 510
Db 1733 ATTCAAGTATTCTTGGCCAAAGTGGCGGTGCTGTATGTGTGAAGGAATAGTGTGCTCGC 1792
Qy 511 LeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysValaGlyAla 530
Db 1793 CTGGTTTATGCTCTAGAGAAAGAGGCCAGGCTATATAACCATCAACAAGAGGCTGGTGCC 1852
Qy 531 MetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeu 550
Db 1853 ATGAATGCACCTGGTCCGTGCTCTGCTCTTATCAATGTGTCATACCTATTGAACTTG 1912
Qy 551 AspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMet 570
Db 1913 GACTGTGATCACTACATCAACAATAGCAAGGCCATAAAGAGGCTATGTGTTTCATGATG 1972
Qy 571 AspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIle 590
Db 1973 GATCTTGGTGGGGAAGAAAGTGTGTATGTACAGTTCCCTCAGAGGTTTGTATGTTAT 2032
Qy 591 AspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGly 610
Db 2033 GACAAATATGATCATGATAGCTTAA CAGGAACGTTGCTTTTGTGATCAACAATGAAGGT 2092
Qy 611 LeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgGlnAla 630
Db 2093 TTGGACGGTATTCAAGGACCCATTTATGTGGGTACTTGGATGTGTTTTCAGACGGCAGCA 2152
Qy 631 LeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAspCys 650
Db 2153 CTGTATGTTATGATGCTCTCTTAAA---ACGAAGAGGCCCAACCATCAAGAACCTTGCAACTGC 2209
Qy 651 CysPro-----CysPheGlyArgLysLysArgLys----- 660
Db 2210 TGGCCCAAGTGGTGCCTCTCTGTCTGTGTCAGCAGGAACAAGAATAAAGAGACTACA 2269
Qy 660 ----- 660
Db 2270 AAACCAAGACGAGAGAAGAAAGATTTATTTTCAAGAAAGCAGAAACCCATCTCCT 2329
Qy 661 ---HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLys 679
Db 2330 GCATATGCTTGGGTGAAATTTGATGAAGGTCTCCA-----GGTGTGATATCGAGAAG 2383
Qy 680 GluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheVal 699
Db 2384 GCCGGAATCGTAATCAACAGAAACTAGAGAAGAAATTTGGGCGAGTCTTCTGTTTCTGTC 2443
Qy 700 ThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeu 719
Db 2444 GCATCAACACTCTTGTGAAACGGAGGACCCCTGAAGAGCGCAAGTCCAGCTTCTCTCTG 2503
Qy 720 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGlu 739
Db 2504 AAGGAAGCTATACATGTTATCAGCTGGCGGTACCAAGACAAGACCCGACTGGGGAAGAG 2563
Qy 740 LeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCys 759
Db 2564 ATTGGCTGGATTTACGATCGATCAGAGAGATATCTTGTGATTTAAGATGCACCTGC 2623
Qy 760 ArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaPro 779
Db 2624 CATGGCTGGCGGTCTATTTACTGTCATCCGAAGCGGCTGCAATTCAAAGGTTCTGCGCT 2683
Qy 780 IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIle 799
Db 2684 CTGAACCTTTCCGACCGCTTTCACCGAGTCTCTCGCTGGGCGCTTGGGTCCGTCGAAAT 2743

```
Qy 800 PhePheSerArgHisSerProLeuLeuTyrGlyTyrIysAsnGlyAsnLeuLysTrpLeu 819
Db 2744 TTCTTTCAGCAAGCACTGCCCACTTTGGTACGGATAC-----GGCGCGGGCTAAAAATTCCTG 2800
Qy 820 GluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAla 839
Db 2801 GAAAGGTTTCTTATATCAACTCATCGTTATCCCTGGAGCTCCATTCCTCTCTCTGGCT 2860
Qy 840 TyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSer 859
Db 2861 TACTGTACCTTGGCTGCCATCTGCCTCTACGGGGAGTTTATCACACGAGCTTACC 2920
Qy 860 ThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeu 879
Db 2921 AATGTCCGCAAGTATCTGGTTCAATGGCACTTTTCATCTGCATCTCCGTCGACCGCATCTG 2980
Qy 880 GluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpVal 899
Db 2981 GAAATGAGGTGGAGTGGCGTGGCCATCGACGACTGGTGGAGGAACGAGCAGTTCCTGGGTC 3040
Qy 900 IleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAla 919
Db 3041 ATCGAGGCGTTTCGGCGCATCTGTTCGGGTGTTCACGGGCTGCTGAAGTGTTCGCC 3100
Qy 920 GlyIleAspThrAsnPheThrValThrSerIysAlaThrGlyAspGluAspGluPhe 939
Db 3101 GGCATCGACAGCACTTCACCGTGCAGCTCGAAGGCC---GGGGAC---GACGAGGAGTTC 3154
Qy 940 AlaGluLeuTyrAlaPheIleTyrThrThrLeuLeuIleProProThrThrLeuLeuIle 959
Db 3155 TCGGAGCTGTACAGTTCAGGTGACCACTGCTGTATACCCCGACACGCTCTCTCTG 3214
Qy 960 IleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSer 979
Db 3215 CTGAACCTTCATCGGGTGTGGCGCGATCTCGAACGCGATCAACAACGGGTACGAGTCG 3274
Qy 980 TrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisLeuTyrPro 999
Db 3275 TGGGGCCCCCTGTTCGGGAAGCTCTTCTCCGCTTCTGGGTGATCGTCCACCTGTACCCG 3334
Qy 1000 PheLeuLysGlyLeuMetGlyValGlnAsnArgThrProThrValValIleTrpSer 1019
Db 3335 TTCTCTAAGGTCTGTGGGGAGGAGCAAGACGACGCGACGATCGTCTGTCTGTCTG 3394
Qy 1020 IleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArg 1039
Db 3395 ATCTGTCTGGCTCGATCTTCTGCTCTGTGGGTCCGCGTCGACCCGCTTCTCGCCAAG 3454
Qy 1040 ThrIysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3455 AGCAACGGCCGCTCTCGGAGGAGTGTGGCCTGGACTGC 3493
```

RESULT 10

```
AAZ99509
ID AAZ99509 standard; DNA; 3813 BP.
XX
AC AAZ99509;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Location/Qualifiers
FT 215..3493
FT CDS /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
FT
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 215 ATGGAGGCTAGCGGGGCTGTGTGGCGGCTCGCATACCGAAGACGAGCTGTGTGTGATC 274
Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnVal----- 36
Db 275 CGC-----CGCGACCGGAGTCCGGAGCCCGCGCGCGCGCGCGCGCGGAG 328
Qy 37 -----CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheVal 54
Db 329 CGCGCGTGCACAGATATGCGGCGAGGAGTCCGGGTGGGCTTCGACGGGAGCCCTTCGTG 388
Qy 55 AlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGlu 74
Db 389 GCGTGCAACGAGTGGCCCTTCCCGCTCGCCCGCGCTCTACGAGTACGAGGCGCGCGAG 448
Qy 75 GlyThrGlnAsnCysProGlnCysLysThrArgTyrIysArgLeuLysGlySerProArg 94
```

```
XX WO200009706-A2.
XX
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US018760.
XX
XX 17-AUG-1998; 98US-0096822P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
XX WPI; 2000-224343/19.
XX P-PSDB; AAY84113.
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
XX
XX Claim 1; Page 129-134; 119pp; English.
XX
XX The present sequence encodes a maize cellulose synthase polypeptide. The
XX cellulose synthase can be used for the improvement of stalk quality for
XX improved stand or silage. It also provides an increased concentration of
XX cellulose in the pericarp, hardening the kernel and improving its
XX handling ability. The sequences are used to produce transgenic plants and
XX seeds expressing the cellulose synthase. The polynucleotide is used for
XX modulating, preferably increasing, the level of the synthase in a plant
XX cell. The plants are preferably monocots. The polynucleotide is also used
XX as a probe or primer in the detection quantitation or isolation of gene
XX transcripts. The probes are useful in detecting deficiencies in the level
XX of mRNA in screenings for desired transgenic plant, for detecting or
XX mutations in the gene, for monitoring upregulation of expression or
XX changes in enzyme activity in screening assays of compounds, for
XX detection of any number of allelic variants of the gene, or for use as
XX molecular markers in plant breeding programs. The isolated nucleic acids
XX of the present invention can also be used for recombinant expression of
XX their encoded polypeptides or for use as immunogens in the preparation
XX and/or screening of antibodies. The proteins can be employed in assays
XX for enzyme agonists or antagonists of enzyme function or for use of
XX immunogens or antigens to obtain antibodies specifically immunoreactive
XX with a protein
XX
XX Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 U; 0 Other;
```

```

449 GACTCGCAAGCGTGGCCGCGAGTGGAGGACCCGCTACAGCGCTCAAGGGCTGCCGCGG 508
Qy ValalaGlyAspAspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAsp 114
Db GTGGCCGCGCAGCAGGAGGAGGAGCGCGCTCGACGACCTGGAGGGCGAGTTCGGCCCTGCAG 568
Qy AspGluAsnGlnGlnArgGlnLeuGluGluArgAsnMetGlnAsnSerGlnIleThrGluAla 134
Db GACGGCGCGCCCAACAGGAC-----GACCCGACGATACGTCGCCGAGTCC 613
Qy MetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGlyAspGly---AsnAsn 153
Db ATGCTCAGGGCCGACATGAGCTACGGCCGCGCGCGCGCGCGCCGCCGCCCTTCAGCCCC 673
Qy ThrProGlnIleProProIleThrGlySerArgSerValProValSerGlyGlu--- 172
Db GTCCCAACAGTCCGCTCCTCACCAACGCGCCAGATGGTTGATGACATCCCGCGGAGCAG 733
Qy -----PhePro-----IleThrAsnGlyTyrGlyHisGlyGluValSerSer 187
Db CACGCGCTCGTCCCTCTACATAGCGCGCGCGCGCGCGCGG----- 778
Qy LeuHisLysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db -----AAGAGATCCACCGCTCCTTCGCGATCCCAACCTTCCAGTGCACCGAGA 832
Qy LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Db TCCATGGACCCGCTCCACAGATCTGGCGCTACCGATATGGCAGCGTGGCCCTGGAGGAG 892
Qy ArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGly 230
Db AGAATGGAGGCGTGAAGCAGAGCAGGAGCGCCTGCAGCATGTTCAGGAGCGAGGCTGCG 952
Qy AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeu 250
Db GGTGATGGGATGGCGACGATGCATCTGCCACTAATGGATGAAGTAGGCGGACCAATG 1012
Qy SerArgLysValSerIleAspSerLysValAsnProTyrArgMetValIleValVal 270
Db TCCAGAAAGTCCCTATATCATCAAGCGAATTAATCCCTACAGCATGATTATCGTTATC 1072
Qy ArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAla 290
Db CGGTTGGTGGTTTGGGTTCTCTTCCACTACCGAGTGCATCCGCGGAAAGATGCA 1132
Qy IleGlyLeuTrpLeuValSerIleCysGluIleTrpPheAlaIleSerTrpIleLeu 310
Db TTGCAITGTGGCTCATATCTGTAATCTGTGAATCTGGTTGGCATGTCTGGATTCTT 1192
Qy AspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeu 330
Db GATCAGTTCCCAAGTGGCTTCCANTCGAGAGAGACTTACCTGGACCGTTGTCACTA 1252
Qy ArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThr 350
Db AGGTTTGAAGAGAGGTCAACCCCTCTCAGCTTGTCTCCAAATCGACTTCTTGTTCAGTACG 1312
Qy ValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAla 370
Db GTTGATCCCAAGAGAACCTCCCTTGGTCACAGGGAACACTGTCTTCCATCTCTTCT 1372
Qy ValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyValAspMetLeu 390
Db GTGGATTATCCGTTGAGAAGTCTCTCTGCTATGTTTCTGATGATGGTGTCTGCAATGCTT 1432
Qy ThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLys 410
Db ACGTTTGAAGCATTTGTCTGAAACATCTCAATTTGCAAGAAATGGGTTCTCTTTCAGCAA 1492
Qy LysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeu 430

```

```
Qy 760 ArgGlyTrpArgSerValTyrCysMetProlyseArgAlaAlaPheLysGlySerAlaPro 779
Db 2624 CATGGCTGGGGCTATATTTACTGCATCCGAGCGGCTCGATTCAAAGGTTCTGGCGCT 2683
Qy 780 IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluile 799
Db 2684 CTGAACCTTTCCGACCGCTTTCACAGGTCCTTCGCTGGGCGCTTGGGTCGTCGAAT 2743
Qy 800 PhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeu 819
Db 2744 TTCTTCAGCAAGCACTGCCACTTTGGTACGGAAC---GGCGGGGCTTAAATTCCTG 2800
Qy 820 GluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAla 839
Db 2801 GAAAGGTTTTCTTATATCAATCCATCGTTCATCCCTGGAGCGTCATTCCTCTCTGGCT 2860
Qy 840 TyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSer 859
Db 2861 TACTGTACCTTGGCTGCATCTGCTGCTCAGGGGAGTTTATCACACAGAGCTTACC 2920
Qy 860 ThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeu 879
Db 2921 AATGTCGCCAGTATCTGTTTCATGCGCACTTTCATCTGCATCTCCGTGACCGCATCTG 2980
Qy 880 GluMetArgTrpSerGlyValSerIleGluTyrTrpArgAsnGluGlnPheTrpVal 899
Db 2981 GAAATGAGTGGAGTGGCGTGGCCATCGAGTGGTGGAGGAAACGAGCAGTTCCTGGGTC 3040
Qy 900 IleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAla 919
Db 3041 ATCGAGCGGTTTCGGGCGATCTGTTTCGGGTGTTCCAGGCGCTCTGAGAGTGTTCGCC 3100
Qy 920 GlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPhe 939
Db 3101 GGCATCGACACGAGCTTCACCGTCAGCTCGAAGGCC---GGGGAC---GACGAGGAGTTC 3154
Qy 940 AlaGluLeuTyrAlaPheLysTrpThrLeuLeuLeuLeuProProThrThrLeuLeuile 959
Db 3155 TCGAGTGTGTACAGTTTCAAGTGAACCACTGCTGATACCCCGACACAGCTCTCTCTG 3214
Qy 960 IleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSer 979
Db 3215 CTGAACCTTCATCGGGTGGTGGCGGATCTCGAAGCGGATCAACACGGGTACGAGTGC 3274
Qy 980 TrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrPro 999
Db 3275 TGGGCGCCCTGTTTCGGGAAGCTCTCTTCGCTCTCTGGTGCATCGTCCACCTGTACCCG 3334
Qy 1000 PheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSer 1019
Db 3335 TTCTCAAGGGTCTGGTGGGAGGCGAGAACAGGAGCGGAGCGATCGTCTGTGGTCC 3394
Qy 1020 IleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArg 1039
Db 3395 ATCTCTGCTGGCTCGATCTTCTCGCTCTCTGGGTCTCGGTCGAGCCCTTCTCGCCAG 3454
Qy 1040 ThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3455 AGCAACGCGCGCTCTCTGGAGGAGTGGCTGGACTGC 3493
```

RESULT 11

AAZ99524
ID AAZ99524 standard; DNA; 3813 BP.

XX AC AAZ99524;

XX DT 03-JUL-2000 (first entry)

XX DE DNA encoding a maize cellulose synthase.

XX KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
transgenic plant; plant breeding marker; ss.

```
XX Zea mays.
XX OS
XX FH Key Location/Qualifiers
XX CDS 215..3493
FT /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
```

XX WO200009706-A2.

XX PD 24-FEB-2000.

XX PF 16-AUG-1999; 99WO-US018760.

XX PR 17-AUG-1998; 98US-0096822P.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX DR WPI; 2000-224343/19.

XX DR P-PSDB; AAY84118.

XX PT New genes which encode maize cellulose synthase polypeptides in plants
useful for modulating the expression of cellulose synthase in plants and
to produce transgenic plants expressing the novel protein.

XX PS Claim 1; Page 168-173; 119pp; English.

XX CC The present sequence encodes a maize cellulose synthase polypeptide. The
cellulose synthase can be used for the improvement of stalk quality for
improved stand or silage. It also provides an increased concentration of
cellulose in the pericarp, hardening the kernel and improving its
handling ability. The sequences are used to produce transgenic plants and
seeds expressing the cellulose synthase. The polynucleotide is used for
modulating, preferably increasing, the level of the synthase in a plant
cell. The plants are preferably monocots. The polynucleotide is also used
as a probe or primer in the detection quantitation or isolation of gene
transcripts. The probes are useful in detecting deficiencies in the level
of mRNA in screenings for desired transgenic plant, for detecting or
mutations in the gene, for monitoring upregulation of expression or
changes in enzyme activity in screening assays of compounds, for
detection of any number of allelic variants of the gene, or for use as
molecular markers in plant breeding programs. The isolated nucleic acids
of the present invention can also be used for recombinant expression of
their encoded polypeptides or for use as immunogens in the preparation
and/or screening of antibodies. The proteins can be employed in assays
for enzyme agonists or antagonists of enzyme function or for use of
immunogens or antigens to obtain antibodies specifically immunoreactive
with a protein

XX SQ Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3813
Score:	3974.50	Matches:	752
Percent Similarity:	78.53%	Conservative:	122
Best Local Similarity:	67.57%	Mismatches:	158
Query Match:	70.37%	Indels:	81
DB:	3	Gaps:	17

US-10-627-132-30 (1-1052) x AAZ99524 (1-3813)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuile 20

Db 215 ATGGAGGCTAGCGCGGCGCTGGTGGCGGCTCGCATTAACCGGAACGAGCTGGTGGTGATC 274

Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnVal----- 36

Db 275 CGC-----CCGACCGCGAGTCGGAGCCCGGGCGCGCGCGCGCGCGGAG 328

Qy 37 -----CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheVal 54

Db 329 GCGCGTGCAGATATCGCGGACAGAGTGGGGTTCGACGGGAGCCCTTCGG 388
QY 55 AlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgGlu 74
Db 389 GCGTGCACAGTGGCGCTTCCTCCGCTCTCGCGCGCTCTACGAGTACGAGCGCGCGAG 448
QY 75 GlyThrGlnAsnCysProGlnCysTyrThrArgTyrGlyArgLeuLysGlySerProArg 94
Db 449 GCGTGCACAGTGGCGCGCTTCCTCCGCTCTCGCGCGCTCTACGAGTACGAGCGCGCGAG 508
QY 95 ValAlaGlyAspAspGluGluAspLeuGluHisGluPheAsnIleAsp 114
Db 509 GTGGCGGCGCAGAGAGAGCGGCTCGACGCTCGAGCGGCGAGTTCGGCTCGAG 568
QY 115 AspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAla 134
Db 569 GACGCGCGCGCCACAGGAGC-----GACCCGCGAGTACGTCGCGAGTCC 613
QY 135 MetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGlyAspGly---AsnAsn 153
Db 614 ATGCTCAGCGCGCGAGTACGCTACGCGCGCGCGCGCGAGCGCACCCCGGCTTCAGCC 673
QY 154 ThrProGlnIleProIleThrGlySerArgSerValProValSerGlyGlu--- 172
Db 674 GTCCCAACAGTCCGCTCTACCAACGCGCGAGATGGTTGATGACATCCCGCGGAGCAG 733
QY 173 -----PhePro-----IleThrAsnGlyTyrGlyHisGlyGluValSerSer 187
Db 734 CACGCGCTCGTCCGCTCTACATGAGCGCGCGCGCGCGGCG----- 778
QY 188 LeuHisLysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Db 779 -----AAGAGGATCCACCGCTCCCTTCGAGATCCCAACCTTCCAGTGCACCGAGA 832
QY 204 LysTrpAspGluLysLysGlu-----ValSerTrpLysGlu 215
Db 833 TCCATGGACCGCTCCAGGATCTGGCGGCTACCGGATATGGCAGCGTGGCGTGGAGGAG 892
QY 216 ArgMetAspAspTrpLysSerLysGlnGlyLeu-----GlyGlyGly 230
Db 893 AGAATGAGGCGGTGGAAGCAGAGCAGGAGCGGCTTCAGCATGTTCAGGCGGAGGTTGCG 952
QY 231 AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeu 250
Db 953 GGTGATCGGATGGCGAGCATGTCAGTCTGCCACTAATGGATGAAGCTAGGAGCAGCATG 1012
QY 251 SerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 270
Db 1013 TCCAGAAAGTCCCTATATCATCAAGCCGAATTAATCCCTCAGGATGATTATCGTTATC 1072
QY 271 ArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAla 290
Db 1073 CGGTGGTGGTGTTCGGTTCCTTCCTCCATCCAGGATGATGATCCCGCGGAGAGATGCA 1132
QY 291 IleGlyLeuTrpLeuValSerIleLeuCysGluIleTrpPheAlaIleSerTrpIleLeu 310
Db 1133 TTTGCAATTTGGCTCATATCTGTAATCTGTGAATCTGGTTCGATGCTCTGGATTCCT 1192
QY 311 AspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeu 330
Db 1193 GATCAGTTCCTCCAAAGTGGCTTCCATCAGAGAGAGACTTACCTGGAGCGGTTTGTCACTA 1252
QY 331 ArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThr 350
Db 1253 AGTTTGACAGGAAGGTCAACCTCTCAGCTTGTCTCAATCGACTCTTTTGTTCAGTAG 1312
QY 351 ValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAla 370
Db 1313 GTTGATCCCAAGGAACCTCCCTTGGTTCACAGCGAACAACGTCTCTTCCATCTCTTCT 1372
QY 371 ValAspTyrProValAspLysValSerCysTyrTrpValSerAspAspGlyAlaSerMetLeu 390

Db 1373 GTGGATTATCCGGTTGAGAAGGCTCTCTGCTATATGTTTCTGATGATGGTGTCTCAATGCTT 1432
QY 391 ThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLys 410
Db 1433 ACGTTTGAGCATCTGCTGAAACATCTGAATTTGCAAGAAATGGGTTCTTCTTCAGCAA 1492
QY 411 LysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeu 430
Db 1493 AAGTTTAATATCGAGCCTCGTCTCTGAGTGGTATCTTCCAAACAGAGATAGACTACCTG 1552
QY 431 LysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyrGlu 450
Db 1553 AAAGACAAGGTTGCTGCTTCATTTGTTAGGAGAGAGGCGCATGAAGAGAGATACGAG 1612
QY 451 GluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGly 470
Db 1613 GAATTCAGGTAAGGATCAATGCTTGTTCGAAAAGCCCAAAAGTTCTCTGAGGAGGA 1672
QY 471 TrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMet 490
Db 1673 TGGACAATGCAAGATGGAAGCCCTGCGCTGGAAAACAACGATACGCGATCATCTCTGAA 1732
QY 491 IleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArg 510
Db 1733 ATTCAAGTATTCCTTGGCCAAAGTGGCGGTCTGATGTGGAAGGAAATGAGTTGCTCGC 1792
QY 511 LeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAla 530
Db 1793 CTGGTTATGCTCGAGAGAAAGAGCGGCTATATAACCATCACAGAGGCTGCTGCTGCC 1852
QY 531 MetAlaAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeu 550
Db 1853 ATGAATGACATGCTCGGTCTCTGCTCTTATCAAAATGCTGCTACCTATTGAACCTG 1912
QY 551 AspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMet 570
Db 1913 GACTGTGATCACTACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATG 1972
QY 571 AspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIle 590
Db 1973 GATCCTTTGGTGGGAGAAAGTGTGTATGTATGATACAGTTCCTTCAGAGTTTGTATGTT 2032
QY 591 AspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGly 610
Db 2033 GACAAAATGATCATGATCGCTTAAACAGGAACCTTCTCTTTTGTGACATCAACATGAAGGT 2092
QY 611 LeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAla 630
Db 2093 TTGACCGGTATTCAGGACCCATTTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCA 2152
QY 631 LeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAspCys 650
Db 2153 CTGTATGTTATGATGCTCCTTAA--ACGAAAGAGCCACCATCAAGAGCTTGCACACTGC 2209
QY 651 CysPro-----CysPheGlyArgLysLysArgLys----- 660
Db 2210 TGGCCCAAGTGGTGGCTCTCTTGTGCTGCGACGAGCAACAGAAATAAAAGAGACTACA 2269
QY 660 ----- 660
Db 2270 AAACCAAGCGGAGAGAGAAAGATATTTTTCAGAAAGCAGAAACCCCATCTCCT 2329
QY 661 ---HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLys 679
Db 2330 GCATATGCTTTGGGTGGAATTTGATGAAGGTGCTCCA-----GGTGTGATATCGAAG 2383
QY 680 GluMetLeuMetSerHisMetAsnPheGlyLysPheGlyGlnSerAlaAlaPheVal 699
Db 2384 GCGGGAATCGTAATCAACAGAAACTAGAGAGAAATTTGGGCGAGCTTCTCTGTTTTCTC 2443
QY 700 ThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeu 719
Db 2444 GCATCAACACTTTTGAACCGGAGGACCCCTGAAGAGCGCAAGTCCAGCTTCTCTCTG 2503

```
Qy 720 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrrpGlyLeuGlu 739
Db 2504 AAGGAGCTATACATGTTATCAGCTGGCGCTACGAGCAAGACCGCTGGGGAAGAAG 2563
Qy 740 LeuGlyTrrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCys 759
Db 2564 ATTGGCTGGATTACGATCGATCGATCACAGAGATATCTTGACTGGATTTAAGATGCATGC 2623
Qy 760 ArgGlyTrrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaPro 779
Db 2624 CATGGCTGGCGGCTATTACTGATCCGAGCGGCTGCATTCAAAGGTTCTGGCGCT 2683
Qy 780 IleAsnLeuSerAspArgLeuAsnGlnValIleuArgTrrpAlaLeuGlySerValGluIle 799
Db 2684 CTGAACCTTTCCGACCGCTCTTCACGAGTCTTCGCTGGGCGCTTGGGTCCGTCGAAT 2743
Qy 800 PhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrrpLeu 819
Db 2744 TTCTTCAGCAAGCACTGCCACATTGGTACGGATAC--GGCGGGCGGCTAAAATTCTCTG 2800
Qy 820 GluArgPheAlaTrrpIleAsnThrIleTyrProPheThrSerLeuProLeuLeuAla 839
Db 2801 GAAAGGTTTCTTATATCACTCATCGTTTATCCCTGGACGTCATTCCTCTCTGGCT 2860
Qy 840 TyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSer 859
Db 2861 TACTGTACCTTGCCTGCATCGCTGCTACGGGGAAGTTTATCACACAGAGCTTACC 2920
Qy 860 ThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeu 879
Db 2921 AATGTCGCAGTATCTGGTTATCGCATCTTTCATCTGCATCTCCGTCGACCGGCATCTG 2980
Qy 880 GluMetArgTrrpSerGlyValSerIleGluIuTrrpArgAsnGluGlnPheTrrpVal 899
Db 2981 GAAATGAGTGGAGTGGCGTGGCCATCGACACTGGTGGAGGAACGAGCACTTCTGGGTC 3040
Qy 900 IleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuValLeuAla 919
Db 3041 ATCGAGGCGTTTCGGCGCATCTGTTCGGGTGTTCCAGGGCCCTGTAAGGTGTTTCGCC 3100
Qy 920 GlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPhe 939
Db 3101 GGCAATCGACACGAGCTTCACCGTCAGCTCGAAGGCC--GGGGAC--GACGAGGAGTTC 3154
Qy 940 AlaGluLeuTyrAlaPheLysTrrpThrLeuLeuIleProProThrThrLeuLeuIle 959
Db 3155 TCGGAGCTGTACAGTTCAGTGGACCACTGCTGTATACCCCGCACCGCTCTCTCTG 3214
Qy 960 IleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSer 979
Db 3215 CTGAACCTTCATCGGGGTGGTGGCGGATCTCGAACCGGATCAACACGGGTACGAGTCG 3274
Qy 980 TrrpGlyProLeuPheGlyLeuPhePheAlaPheTrrpValIleValHisLeuTyrPro 999
Db 3275 TGGGGCCCCCTGTTCCGGAAGCTCTTCTCGCTTCTGGGTGATCGTCCACCTGTACCCG 3334
Qy 1000 PheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrrpSer 1019
Db 3335 TTCTCTCAAGGCTCTGGTGGGAGGAGCAACAGGACCGCGACGATCGTCATCGTCTGTGTC 3394
Qy 1020 IleLeuLeuAlaSerIlePheSerLeuLeuTrrpValArgIleAspProPheIleValArg 1039
Db 3395 ATCTGCTGGGCTCGATCTTCGCTCTGGGTGCGGCTCGACCCCGTTTCTCGCCAAAG 3454
Qy 1040 ThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3455 AGCAACGGCCCGCTCTCTGGAGGAGTGGCTGGACTGC 3493
RESULT 12
AAC49550
ID AAC49550 standard; DNA; 3198 BP.
XX
```

```
AC AAC49550;
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 61572.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
```


133 CCAGTTTGTGGCGCTTCTAGAGTATGAGAGGAAAGATGGAATCAATCTTGTCTCTCAG 192
Qy
82 CyalybThrArgTyrLysArgLeuLysGlySerProArgValalaglyAspAspGlu 101
Db
193 TGCAAAACCCAGATACAAAGAGCTCAAGAGTGTCTCTATTCCTGGTGATAAAGACGAG 252
Qy
102 GluAspIle--AspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db
253 GATGGCTTGTGATGAGTGTGTTGAGTTCACTACCTT----- 294
Qy
121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db
295 -----CAGAAGCAGAAAATTTCCAGACGGATCTTGTGTGGCATCTT 336
Qy
141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro----- 158
Db
337 ACTCGTGGGAAGGAGAGAAATGGGGAAACCCAGTATGATAAAGAGGTCTCTCACAAAT 396
Qy
159 -----ProIleIleThrGlySerArgSerValProValSerGlyGluPhe----- 173
Db
397 CATCTTCTGTCTCACGAGCAGACAAGAT-----ACTTCAGGAGAGTTTCTGTGCTGCC 450
Qy
174 -----ProIleThrAsnGly-----TyrGly 180
Db
451 TCACCTGAACGCTCTCTGTATCTTCTACTATCGCTGGGGAAAGCGCTTCCCTATTCA 510
Qy
181 HisGlyGluValSerSerLeuHisLysArgIleHisProTyrProValSerGluPro 200
Db
511 ---TCAGATGTCAATCAATCAACCAATAGAAAGGAT-----GTGGATCTCTGTT 555
Qy
201 GlySerAlaLysTyrAspGluLysGluValSerTrpLysGluArgMetAspAspTrp 220
Db
556 GGACTCGGG-----AATGTAGCTTGGAAAGGAGAGTGTGTGCTGG 597
Qy
221 LysSerLysGln-----GlyLeuGly-----Gly 229
Db
598 AABATGAAGCAAGAGAATACTGGTCTCTCAGCAGCGAGGCTGCTTCGAAAGAGGT 657
Qy
230 GlyAlaAspProGlu-----AspMetAspAlaAspValAla--LeuAsnAspGlu 245
Db
658 GGAGTAGATATTGATGCCAGCACAGATATCTCAGCAGATGAGGCTCTGCTGAATGACGAA 717
Qy
246 AlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArg 265
Db
718 CGAGGCGAGCTCTGTCAAGAAAGTTTCAATTCCTTCATCCGATCAATCTTACAGA 777
Qy
266 MetValIleValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHis 285
Db
778 ATGGTTATTATGCTGCGGCTTGTATTCTCTTGTCTCTTCTTGTGCATTACCGTATAACAAAC 837
Qy
286 ProValProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAla 305
Db
838 CCAGTGCCTAATGCTTGTCTATGGCTGTCTGTGTATATGATGTGATCTGCTGTTGCC 897
Qy
306 IleSerTrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeu 325
Db
898 TTATCTCGATTTTGGATCAGTTTCCCAAGTGGTTTCTGTGACCGGTGAACCTACCTC 957
Qy
326 AspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAsp 345
Db
958 GACAGGCTTGTCTTTAAGATATGATCGTGAAGGTGAGCCATCACAGTTAGCTGCTGTTGAC 1017
Qy
346 LeuPheValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrVal 365
Db
1018 ATTTTCGTGAGTACTGTTGACCCCTTGAAGGAGCCACCCTTGTGACAGCCCAACACAGTG 1077
Qy
366 LeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAsp 385
Db
1078 CTCCTATTCTGGCTTTGATCATTACCCAGTTGACAAGGTGCTCTGTATTATGTTCTGATGAT 1137
Qy
386 GlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrp 405
Db
1138 GGTGCTGCTATTGTTATCTTGAATCACTTGTGAGAAAACATCAGAGTTTGTGCTGTAATGG 1197

Qy
406 ValProPheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeu 425
Db
1198 GTACCAATTTTGAAGAATAATACCATAGAGCTCGTGCCACCAGAAATGTTCTGCTGG 1257
Qy
426 LysValAspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMet 445
Db
1258 AAAATAGATTACTTGAAGCATAAAGTTCCAGACATCTTTGTCAAAGATCCTAGAGCTATG 1317
Qy
446 LysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLys 465
Db
1318 AAGAGGGAATATGAGGAATTTAAATCCGAATCCAGATCTTGTTCCTCAAAGCCCTAAAA 1377
Qy
466 ValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArg 485
Db
1378 TGTCTCGAAGAGGGTGGGTATGCAAGATGGCACCGCTGGCTGGNAATATATACAAGG 1437
Qy
486 AspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGly 505
Db
1438 GACCATCCAGGAATGATCCAGGTCTTCTTAGGGCAAAATGTTGACTTGTATGCAGAGGC 1497
Qy
506 AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHis 525
Db
1498 AATGAGCTCCCGGTTTGGTATATGTTCTCGAAGAAAGCGACAGATTCACAGCCAC 1557
Qy
526 LysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaPro 545
Db
1558 AAAAGGCTGCTGTATGATGACCTGGTGAGAGTTTCAGCAGTCTTCTACCAATGGACCT 1617
Qy
546 PheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAla 565
Db
1618 TTCATCTTGAATCTTGATTTGATGATCATACATAATAACAGCAAAAGCCTTAAAGAGAACA 1677
Qy
566 MetCysPheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGln 585
Db
1678 ATGTGCTTCTGTAGTGGACCCAAACCTCGGGAAGCAAGTTTGTATGTTTCAGTTCCTCCACA 1737
Qy
586 ArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAsp 605
Db
1738 AGATTGTATGATTCGATAGAACAGATAGATATGCTAATCGTAATACCGTGTCTTTGAT 1797
Qy
606 IleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysVal 625
Db
1798 ATTAACCTTGAGAGGTTTAGATGGATTCAGAGCCTGTATATGTCGGAACCTGATGTGTT 1857
Qy
626 PheArgGlnAlaLeuTyrGlyTyrAsnProPro-----LysGlyProLysArgPro 643
Db
1858 TTCACACAGACAGCATTTATACGGTTATGAACCTCCCAATAAAAGTAAACACACAGAGCCA 1917
Qy
644 LysMetValThrCysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLys 663
Db
1918 AGTCTTTTATCTAAGCTCTGTGGTGA-----TCAAGAAAGAAGAAATTCCAAAGCTAAG 1971
Qy
663 ----- 663
Db
1972 AAAGAGTCGGCAAAAAGAAATCAGCGAGGCATACTGACTCAACTGTTCTCTGTATTCAAC 2031
Qy
664 ---AspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGluMetLeu 682
Db
2032 CTCGATCAGCATAGAAAGGAGGTTGAAGGTGCTGCTGTTTGTATGATGATAAAGCGCTCTTA 2091
Qy
683 MetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThrSerThr 702
Db
2092 ATGTCCGAATGAGCTCGGAGACCGATTTGGACAGTCTGCTGTTTGTGTTCTTCTTACC 2151
Qy
703 LeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuLysGluAla 722
Db
2152 CTAATGCAAAATGGTGTGTTCTCTCTTCAGCAACTCCAGAAAAACCTTCTCAAAGAGGCT 2211
Qy
723 IleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTrp 742
Db
2212 ATCCATGTCTATTGTTGTTGATGAGATTAAGTCAGATTGGGAATGGAGATTGGATGG 2271

```
QY 743 IleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyr 762
Db 2272 ATCTATGGTCTGTGACAGAAATATCTGATGGGTTCAAAATGCATGCCCGGTGATGG 2331
QY 763 ArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIleLeuLeu 782
Db 2332 CGATCCATTACTGTCATGCTTAAGCTTCCAGCTTTCAGGGTTCGCTCATCATCTT 2391
QY 783 SerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIlePheSer 802
Db 2392 TCAGATCGTCTGAACCAAGTGTGAGGTGGGCTTATAGTTTCAGTTGAGATTCCTCTCAGT 2451
QY 803 ArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTyrPheLysArgPhe 822
Db 2452 CGGCATTGCTTATATGGTATGTTAC---AATGGAGGCTAAATTTCTTCAGAGGTTT 2508
QY 823 AlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThr 842
Db 2509 GCGTATGTGAACACCAACCATCTACCTATCCTCCATCTCTCTCATGTATGTATG 2568
QY 843 LeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThrPheAla 862
Db 2569 TTGCCAGCGGTTGTCTCTTCCACCAACCAAGTTTATTTCTCTCAGATTAGTAACATTGCA 2628
QY 863 SerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArg 882
Db 2629 AGTATATGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2688
QY 883 TrpSerGlyValSerIleGluLeuTyrTrpArgAsnGluGlnPheTyrValIleGlyGly 902
Db 2689 TGGAGTGGCTAGGATGACGAAATGGTGGAGAAACGAGCAGTTTGGTCTATTGGTGA 2748
QY 903 ValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAsp 922
Db 2749 GTATCCGCTCATTTATTCGTCTGTTTCAAGGTATCTCAAGTCTCTCAAGTCTCTCAAGT 2808
QY 923 ThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAspGluPheAlaGluLeu 942
Db 2809 ACAAACTTTCACAGTTTACCTTCAAAAGCTTCA---GATGAAGACGGAGACTTTGCTGAGCTC 2865
QY 943 TyrAlaPheLysTyrThrThrLeuLeuIleProProThrThrLeuLeuIleIleAsnVal 962
Db 2866 TACTTGTTCATATGACACACTTCTGATTCGCCCAAGCAGCTCTCTCATTTGTAACCTTA 2925
QY 963 IleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyPro 982
Db 2926 GTGGAGTGTGTCAGGAGTCTCTTATGCTATCAACAGTGGATACCAATCATGGGACCA 2985
QY 983 LeuPheGlyLysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLys 1002
Db 2986 CTCCTTGTGAAGTTGTTCTTTGCCCTTCTGGGTGATTTGTCATTGTACCTTTCTCCTCAAG 3045
QY 1003 GlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTyrSerIleLeuLeu 1022
Db 3046 GGTTCATGGGTGCGACAGACCGGACTCTACCATTTGTGTGTGTGTGTGTGTGTGTGTGT 3105
QY 1023 AlaSerIlePheSerLeuLeuTyrValArgIleAspProPheIleValArgThrLysGly 1042
Db 3106 GCTTCTATCTCTCTGTTGTGGGTAGGATTTGATTCCTTCACTAGCGGAGTCACTGGC 3165
QY 1043 ProAspValArgGlnCysGlyIleAsnCys 1052
Db 3166 CCGGACATTCTGGAATGTGAATCAACTGT 3195
```

RESULT 13

```
AAS16458
ID AAS16458 standard; cDNA; 3799 BP.
XX
AC AAS16458;
XX
DT 14-FEB-2002 (first entry)
DE Corn cDNA encoding cellulose synthase Cqrae19/cesa-19.

XX
KW Corn; ss; cellulose synthase; Cdpgs45; cesa-3; Cqrae19; cesa-9;
  stalk quality; improved stand; silage; pericarp; kernel hardening;
  handling ability; transgenic plant.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 238..3799
FT /tag= a
FT /product= "Cellulose synthase"
XX
PN WO200179516-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US011951.
XX
PR 14-APR-2000; 2000US-00550483.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG;
XX
XX WPI; 2002-041338/05.
DR P-PSDB; AAU10496.
XX
PT New cellulose synthase polypeptides and polynucleotides, useful in
  improving stalk quality or silage, and in increasing concentration of
  cellulose in the pericarp, hardening the kernel for improved handling
  ability.
XX
PS Claim 4; Page 80-85; 89pp; English.
XX
CC The invention relates to isolated nucleic acids encoding two cellulose
  synthase proteins from corn, Cdpgs45 (cesa-3) and Cqrae19 (cesa-9). Also
  disclosed are a recombinant expression cassette comprising the
  polynucleotide (operably linked to a promoter) a host cell comprising the
  recombinant expression cassette and a transgenic plant comprising the
  recombinant expression cassette. The nucleic acid is useful in the
  improvement of stalk quality for improved stand or silage, and in the
  increased concentration of cellulose in the pericarp, hardening the
  kernel, and thus improving its handling ability. The nucleic acids may
  also be used as probes or amplification primers in the detection,
  quantification or isolation of gene transcripts, as probes in detecting
  deficiencies in the level of mRNA, for detecting gene mutations or
  allelic variants, for monitoring up regulation of expression or changes
  in enzyme activity in screening assays, for site directed mutagenesis,
  and in sense or antisense suppression of one or more genes in a host
  cell, tissue or plant. The polypeptides may be used in assays for enzyme
  agonists or antagonists, and as immunogen or antigen to obtain antibodies
  specifically immunoreactive with the protein. The present sequence
  encodes a corn cellulose synthase of the invention
XX
SQ Sequence 3799 BP; 925 A; 931 C; 995 G; 944 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 0 Length: 3799
Score: 3914.50 Matches: 751
Percent Similarity: 78.32% Conservative: 98
Best Local Similarity: 69.28% Mismatches: 150
Query Match: 69.31% Indels: 85
DB: 6 Gaps: 16

US-10-627-132-30 (1-1052) x AAS16458 (1-3799)
QY 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53
Db 283 GGGCAGGTGTGCCAGATCTCGGGCGATGGCGTGGGCACTACGGCGGAGAGCGTCTTC 342
QY 54 ValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArg 73
Db 343 ACCGCCCTGCGAGCTGTGGGGTTCCTCGGTTGCCGCCCTCTCTAGAGTACGAGCGCAAG 402
```

Qy	74	GlulGlyThrGlnAenCysProGlnCysLysThrArgTyrIleArgLeuLysGlySerPro	93
Db	403	GAGGCGACACACGGCTGCCCGAGTGCNAAAAACAAGTACAGCGCCACAGGGAGGTCCA	462
Qy	94	ArgValAlaGlyAspAsp---AspGluGluuAspIleAspAspLeuGluHisGluPheAen	112
Db	463	GCATCCGAGGGGAGGAGGACGATCTGTGTCGCGATGTGTAGC---GACTTTCAC	519
Qy	113	Ile-----AspAspGluAenGlnArgGlnLeuGluGlyAenMetGln	127
Db	520	TACCCTGCATCTGGCAATGACGACCGAAGCAG-----	552
Qy	128	AsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAsp	147
Db	553	-----AAGATGTGTGACAGGATCGCGAGCTGGCGCATGAATGCTGGCGGACGCGGGAT	606
Qy	148	AspGly-----AspGlyAenAen	153
Db	607	GTTGGCGCGCCCCAAGTATGACAGTGTGTGAGATCGGGCTTACCAAGTACGACAGTGTGTGAG	666
Qy	154	ThrProGln---IleProProIleIleThrGlySerArgSerValProValSerGlyGlu	172
Db	667	ATCCCTCGGGATACATCCCGTCAGTCACATAACAGCCAG-----ATTTCGGGAGAA	717
Qy	173	PheProIleThrAenGlyTyrGlyHisGlyGluValSerSerSerLeuHisIysArgIle	192
Db	718	ATCCCTGGTGTCTCCCTGACCATCATATGATGTCTCTACTCGGAAACATTTGGCAGCGC	777
Qy	193	HisProTyrProVal-----SerGluProGlySerAlaIysTrpAspGluLys	208
Db	778	GCCCCATTTCCCTATATGATGATCATTCATCAATCCGTCGAGGGGAATTCCTGTGTACGCTT	837
Qy	209	LysGluValSerTrpLysGluuArgMetAspAspTrpLysSerLysGln-----GlyIle	226
Db	838	GGGAATGTGGCTCGAAGAGAGGGTTGATGGCTGGAAAAATGAAGCAGGACAAAGGAAACA	897
Qy	227	Leu-----GlyGlyGlyAlaAspProGluuAspMet	236
Db	898	ATTCCTCATGACGAATGGCACAGCATTTGCTCCCTCTGAGGGCGGGGTGTGTGTGATATT	957
Qy	237	AspAla-----AspValAlaLeuAsnAspGluAlaArgGlnPro	249
Db	958	GATGCATCAACTGATTACAAACATGGAAGATGCCCTTATTAACCATGAACTGCCAGCCT	1017
Qy	250	LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal	269
Db	1018	CTATCTAGAAAGTTCCCATCTTCCTCTCCAGGATAAATCCATACAGGATGGTCATTGTG	1077
Qy	270	ValArgLeuValValLeuAlaPheLeuArgTyrArgIleLeuHisProValProAsp	289
Db	1078	CTACATTTGATTGTTCTTAGCATCTTTCTGCACTACCGGATCACAAAATCCTGTGCGTAA	1137
Qy	290	AlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIle	309
Db	1138	GCATACCCACTGTGGCTTCTATCTGTTATATGTGAGATCTGGTTGCTCTTCTCTCGATA	1197
Qy	310	LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrIleuAspArgLeuSer	329
Db	1198	TTGGATCATGTTCCAAAGTGGTTCCAAATCAACCGCGAGACTTACCTTGTATAGATCGCA	1257
Qy	330	LeuArgTyrGluuArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer	349
Db	1258	TTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTTTGTTCAGT	1317
Qy	350	ThrValAspProLeuLysGluProProLeuValThrAlaAenThrValLeuSerIleLeu	369
Db	1318	ACTGTGCAACCAATGAAGGAGCGCTCCTCTGTCTCATCTGCCAATACCGTGTATCCATTCTC	1377
Qy	370	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyValaserMet	389
Db	1378	GCTGTGGACTATCTCTGTGTGAATAAGTCTCTTGGCTATGTATCTGTATGATGAGCTGCTATG	1437

390	LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys	409
1438	CTGACATTTGATGAGCTAGCTAGACATTCAGAGATTTGCTAGAAAAATGGGTGGCCATTTGTT	1497
410	LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr	429
1498	AAGAAGTACAACATTAACCTTAGAGTCTCTGAATGGTACTTTCTCCAGAAAAATTGATTAC	1557
430	LeuLysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyr	449
1558	TTGAGGACAAAGTCACCTTCATTGTTAAAGACCGCGGGCCATGAAGAGAGAATAT	1617
450	GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu	459
1618	GAAGAATTCAAAATTAGGGTAAATGGCTTTGTTAAAGGCACAAAAAGTCCCTCAGGAA	1677
470	GlyTyrIleMetLysAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly	489
1678	GGATGGATCATGCAAGATGGCAACCATGGCCAGGAAAAAATACCAAGGACCATCTCTGGA	1737
490	MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro	509
1738	ATGATTTCAGGTTTTCTTGGTTCACAGTGGTGGTCTTGATATCTGAGGTTATGAGCTACCC	1797
510	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysLysAlaGly	529
1798	CGTTTGGTCTATGTTCTTCGTGAAAAACGTCCTGGATTCACGATCACAAGAAAGCTGGT	1857
530	AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn	549
1858	GCCATGAATGCTCTTGTGCGCGTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAAT	1917
550	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu	569
1918	CTTGATTGTGATCATCATCATCAACAACAGTAAAGCTCTCAGGAGAGCTATGTGCTTCCTT	1977
570	MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly	589
1978	ATGGATCCTTAACCTAGGAAGAGTGTCTGCTATGTTTCAGTTTCCCAGAGGTTTCGATGGT	2037
590	IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys	609
2038	ATTGATAGGAATGATCATGATATGCCACAGAAACACCGTGTGTTTCGATATTAACCTCAGA	2097
610	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln	629
2098	GGTCTTGATGGCATCCAAGACCCAGTTTATGTGGGCATCTGGCTGTGTTTCAACAGAAACA	2157
630	AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp	649
2158	GCTCTATATGTTATGATGCCCCCAATTAAGCMAAAGAAGGTGGTGTCTTCTGTCTCA----	2211
650	CysCysProCysPheGlyArgLysLys-----	658
2212	---TCACTATGTGTTATGAGCCCCCAATTAAGCMAAAGAAGGTGGTGTCTTCTGTCTCA----	2266
659	-----ArgLysHisAlaLysAspGlyLeuPro-----	668
2269	AAGTCACAGAAGCATGTGGACAGTCTGTGCGCATATTCATCTTGAAGATATAGAGGAG	2328
669	GlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPhe	688
2329	GGAGTTGAAGCGCTGGATTGATGATGAGAAAAATCATTCTTATGTTCTCAATGAGCTTG	2388
689	GluLysArgPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGly	708
2389	GAGAAGAGATTGGCCAAATCTGCAAGTCTTGTGCGTCCACTCTGATGAGAAATATGGTGGT	2448
709	ValProProSerSerProAlaLeuLeuLysGluAlaIleHisValIleSerCys	728
2449	GTTCCTCAGTCTGGACATCCAGAAATCTCTTCTGAAGAAGAGCTATCCATGCTCATAGTTGT	2508
729	GlyTyrGluAspLysThrAspTrpGlyLeuGluLeuGlyTyrTrpIleTyrGlySerIleThr	748


```
Db 2509 GCTACGAGACAGATTGAATGGGAACTGAGATGGGTGATCTATGTTCTGGACG 2568
Qy 749 GluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMet 768
Db 2569 GAAGATATTTCTCACTGGGTTCAAGATGACGACGAGGCTGGCGGTGATCTACTGATG 2628
Qy 769 ProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGln 788
Db 2629 CCTAAGCGCGCGCTTCAAGGGATCGGCTCCATCAATCTCTCAGACCGTCTGAACACG 2688
Qy 789 ValLeuArgTrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeu 808
Db 2689 GTCTCCGGTGGCTCTCGTTTCAGTGGAAATCTTTTCAGCGGATGTCCTCCCTATGG 2748
Qy 809 TyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThr 828
Db 2749 TACGGGTAC--GGAGGACGCTGAAGTTCTTGGAGAGATTCGCTACATCAACACACC 2805
Qy 829 IleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeu 848
Db 2806 ATCTACCGCTCACGCTCCCGCTCTCTCATTTACTGTATCTGCTGCTGCTGCTGCTG 2865
Qy 849 LeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAla 868
Db 2866 CTCACGGGAAGTTTCATCATCCAGAGATCAGAACTTCGCTAGTATCTGGTTCATCTCT 2925
Qy 869 LeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIle 888
Db 2926 CTCTTCATCTCGATCTCGCCACGGGTATCTCGAGATGAGTGGAGCGGCTGGGCATC 2985
Qy 889 GluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyValSerAlaHisLeuPhe 908
Db 2986 GACGAGTGGTGGAGAACGAGCAGTCTCGGTTCATCGGAGGATCTCCGCCACCTCTTC 3045
Qy 909 AlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr 928
Db 3046 GCGGTCTTCAGGGCTCTCTCAAGGTGCTTGGCGGATCGACCACTTCACCGTCACC 3105
Qy 929 SerLysAlaThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThr 948
Db 3106 TCCAGGGCTCG--GATGAGACGGCGACTTCGGGAGCTGTATCTGTCAGTGGACG 3162
Qy 949 ThrLeuLeuIleProProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGly 968
Db 3163 ACATTTCTGATCCCGCCACCACTCTCTGATCATCAACTGGTGGCGGTGTTGTCGCGC 3222
Qy 969 IleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhe 988
Db 3223 ATCTCTACGCCATCAACAGCGGGTACCAGTCTGGGGTCCGCTCTTCGGCAAGCTCTTC 3282
Qy 989 PheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGln 1008
Db 3283 TTCGCTCTTGGTGTATGTTCACTGTACCGTTCCTCAAGGTCTCATGGTGGCAG 3342
Qy 1009 AsnArgThrProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeu 1028
Db 3343 AACCCAGCCCGCACCATCGTGGTGTCTCTGGGCGATCTCTGCTGGCTGATCTCTCTCT 3402
Qy 1029 LeuTrpValArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCys 1048
Db 3403 CTGTGGGTTCGATCGATCCGTTTACCACCGCGTCACTGGCGCCGATCTCGAACGTGT 3462
Qy 1049 GlyIleAsnCys 1052
Db 3463 GGCATCACTGC 3474
```

RESULT 14

AAZ99512

ID AAZ99512 standard; DNA; 3746 BP.

XX

AC AAZ99512;

XX

DT 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

XX transgenic plant; plant breeding marker; ss.

KW Zea mays.

XX Location/Qualifiers

PH 321..3548

FT /*tag= a

FT /product= "cellulose synthase"

FT /transl_except= (pos: 1800..1802, aa: Xaa)

FT /note= "no termination codon given; Xaa is an unspecified

XX amino acid"

PN WO200009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US018760.

XX 17-AUG-1998; 98US-0096822P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI, 2000-224343/19.

DR P-PSDB; AAY84114.

XX New genes which encode maize cellulose synthase polypeptides in plants

PT useful for modulating the expression of cellulose synthase in plants and

PT to produce transgenic plants expressing the novel protein.

XX Claim 1; Page 137-141; 119pp; English.

CC The present sequence encodes a maize cellulose synthase polypeptide. The

CC cellulose synthase can be used for the improvement of stalk quality for

CC improved stand or silage. It also provides an increased concentration of

CC cellulose in the pericarp, hardening the kernel and improving its

CC handling ability. The sequences are used to produce transgenic plants and

CC seeds expressing the cellulose synthase. The polynucleotide is used for

CC modulating, preferably increasing, the level of the synthase in a plant

CC cell. The plants are preferably monocots. The polynucleotide is also used

CC as a probe or primer in the detection quantitation or isolation of gene

CC transcripts. The probes are useful in detecting deficiencies in the level

CC of mRNA in screenings for desired transgenic plant, for detecting or

CC mutations in the gene, for monitoring upregulation of expression or

CC changes in enzyme activity in screening assays of compounds, for

CC detection of any number of allelic variants of the gene, or for use as

CC molecular markers in plant breeding programs. The isolated nucleic acids

CC of the present invention can also be used for recombinant expression of

CC their encoded polypeptides or for use as immunogens in the preparation

CC and/or screening of antibodies. The proteins can be employed in assays

CC for enzyme agonists or antagonists of enzyme function or for use of

CC immunogens or antigens to obtain antibodies specifically immunoreactive

CC with a protein

XX Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;

SQ

Alignment Scores:

Pred. No.: 0 Length: 3746

Score: 3912.50 Matches: 750

Percent Similarity: 78.96% Conservative: 102

Best Local Similarity: 69.51% Mismatches: 150

Query Match: 69.27% Indels: 77

DB: 3 Gaps: 17

US-10-627-132-30 (1-1052) x AAZ99512 (1-3746)

Qy 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53

[illegible]

QY 714 SerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLys 733
 Db : : : : :
 2541 ACTCCGAGTCTCTTCTGAAGAAGCTATCCATGTTATAGCTGTGGCTATGAGACAAG 2600

QY 734 ThrAspTrpGlyLeuGlyLeuGlyTyrPheLeuGlySerLeuGlyLeuGlyLeuThr 753
 Db : : : : :
 2601 ACTGAATGGGAAGTCTGAGTCTGGGTGATCTACCGTCTCTGTGACAGAAGACATTCACCC 2660

QY 754 GlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAla 773
 Db : : : : :
 2661 GGATTCAAGATGCACGCGGAGGCTGGCGTGCATCTACTGATGCCCAAGCGCCAGCT 2720

QY 774 PheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAla 793
 Db : : : : :
 2721 TTCAAGGGTCTGCCCGCATCAATCTTCCGACCGTCTGAACACGAGTGTCCGGTGGCT 2780

QY 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuGlyTyrGlyLysAsn 813
 Db : : : : :
 2781 CTTGGTCTCGGAGATCTCTTCACCGGACATGCCCTCTGGTACGGCTAC---GGA 2837

QY 814 GlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThr 833
 Db : : : : :
 2838 GGGCGGCTCAAGTCTCTGAGAGATTCGCGTACATCAACACCATCTACCGCTCACG 2897

QY 834 SerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
 Db : : : : :
 2898 TCCATCCCGCTTCTCATCTACTGCTACCTCCCGCCATCTGTCTGCTACCGGAAAGTTC 2957

QY 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
 Db : : : : :
 2958 ATCATCTCAGATCAGACATCTCCGACATCTGGTTCATCTCCCTCTTCATCTCGATC 3017

QY 874 PheAlaThrGlyLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArg 893
 Db : : : : :
 3018 TTCGCCACGGCATCTCGAGATGAGTGGAGGGTGGGCATCGACAGTGTGGGAGG 3077

QY 894 AsnGluGlnPheTrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGly 913
 Db : : : : :
 3078 AACGAGCAGTCTGGGTGATCGGGGCATCTCCGCGCATCTCTCGCGTGTTCAGGCG 3137

QY 914 LeuLeuValValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
 Db : : : : :
 3138 CTGCTCAAGGTCTGCCCGCATCGACACCACTTCACCGTCACTCCAGGCGCTCG--- 3194

QY 934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIlePro 953
 Db : : : : :
 3195 GACGAGGACGGCACTTCGGGAGCTGTACATGTTCAAGTGACGACGCTCTGTATCCG 3254

QY 954 ProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyLysSerAspAlaIle 973
 Db : : : : :
 3255 CCCACCCATCTCTATCATCAACTGTCGGCGTCTCGCGGATCTCTTACGCCATC 3314

QY 974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
 Db : : : : :
 3315 AACAGCGGATACAGTCGCGGCGCGCTCTTCGCGAAGCTCTTCTGCGCTTCGGGTC 3374

QY 994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyValArgGlnAsnArgThrProThr 1013
 Db : : : : :
 3375 ATCGTCCACTGTACCGTCTCTCAAGGCGCTCATGGGCGAGCAGACCGGACCCCGACC 3434

QY 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
 Db : : : : :
 3435 ATCGTCTGCTGTGGGCCATCTCTGTCGGGTCCATCTCTCTGTCGTGGGTTCGCATC 3494

QY 1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db : : : : :
 3495 GACCCCTTACCACCGCGTCTACTGGCCCGGATACCCAGACGTGTGGCATCAACTGC 3551

RESULT 15
 AAZ99527
 ID AAZ99527 standard; DNA; 3746 BP.
 XX

AC AAZ99527;
 XX 03-JUL-2000 (first entry)
 XX DNA encoding a maize cellulose synthase.
 XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW transgenic plant; plant breeding marker; ss.
 XX Zea mays.
 XX Key Location/Qualifiers
 FH 321..3449
 FT /tag= a
 FT /product= "cellulose synthase"
 FT /transl_except= (pos: 1800..1802, aa: Xaa)
 FT /note= "no termination codon given; Xaa is an unspecified
 FT amino acid"
 XX WO200009706-A2.
 PN 24-FEB-2000.
 PD 16-AUG-1999; 99WO-US018760.
 PF 17-AUG-1998; 98US-0096822P.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 PI WPI; 2000-224343/19.
 XX P-PSDB; AAY84119.
 DR New genes which encode maize cellulose synthase polypeptides in plants
 XX useful for modulating the expression of cellulose synthase in plants and
 XX to produce transgenic plants expressing the novel protein.
 XX Claim 1, Page 176-181; 119pp; English.
 XX The present sequence encodes a maize cellulose synthase polypeptide. The
 XX cellulose synthase can be used for the improvement of stalk quality for
 XX improved stand or silage. It also provides an increased concentration of
 XX cellulose in the pericarp, hardening the kernel and improving its
 XX handling ability. The sequences are used to produce transgenic plants and
 XX seeds expressing the cellulose synthase. The polynucleotide is used for
 XX modulating, preferably increasing, the level of the synthase in a plant
 XX cell. The plants are preferably monocots. The polynucleotide is also used
 XX as a probe or primer in the detection quantitation or isolation of gene
 XX transcripts. The probes are useful in detecting deficiencies in the level
 XX of mRNA in screenings for desired transgenic plant, for detecting or
 XX mutations in the gene, for monitoring upregulation of expression or
 XX changes in enzyme activity in screening assays of compounds, for
 XX detection of any number of allelic variants of the gene, or for use as
 XX molecular markers in plant breeding programs. The isolated nucleic acids
 XX of the present invention can also be used for recombinant expression of
 XX their encoded polypeptides or for use as immunogens in the preparation
 XX and/or screening of antibodies. The proteins can be employed in assays
 XX for enzyme agonists or antagonists of enzyme function or for use of
 XX immunogens or antigens to obtain antibodies specifically immunoreactive
 XX with a protein
 XX SQ Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 0 Length: 3746
 Score: 3912.50 Matches: 750
 Percent Similarity: 78.96% Conservative: 102
 Best Local Similarity: 69.51% Mismatches: 150
 Query Match: 69.27% Indels: 77
 DB: 3 Gaps: 17

US-10-627-132-30 (1-1052) x AAZ99527 (1-3746)

QY 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53
Db 366 GGACAGGTGTGCCAGATCTCGCGGACCGCGTGGCCACCGCGGAGGGGAGCTCTTC 425
QY 54 ValAlaCysAenGluCysGlyPheProValCysArgProCysTyrGluTyrGluArg 73
Db 426 CCCGCTCGCAGCTCGCGGTTCGCGGTGCGCGCTCTACGAGTACGAGCGCAAG 485
QY 74 GluGlyThrGlnAenCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerPro 93
Db 486 GACGCGACGCGAGCGTGCCTCCAGTGCAGCAAGTACAAAGCGCCACAAGGGGAGCGCG 545
QY 94 ArgValAlaGlyAspAspGluClnAspIleAspLeuGluHisGluPheAen--- 112
Db 546 CGCATCCGTGGG-----GAGAGGAGCAGCAGCTGATGCCGATGAGCTTCAATTAC 599
QY 113 IleAspAspGluAenGlnArgGlnLeuGluGlyAenMetGlnAenSerGlnIleThr 132
Db 600 CTTCATCTGGCAATGAGGAC-----CAGAAGCAGAAGATTGCC 638
QY 133 GluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly----- 149
Db 639 GACGAATGCGCAGCTGCGCGATGAACGTTGGGGCAGCGGGGATGTTGGTGCGCCCAAG 698
QY 150 -----AspGlyAenAenThrProGln---Ile 157
Db 699 TATGACATGGGAGATCGCGCTTACCAAGTATGACATGCGGAGATTCTTCGGGGATAC 758
QY 158 ProProIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAen 177
Db 759 ATCCATCAGTCACTAACAGCCAG-----ATCTCAGGAGAAATCCCTGGTGTCTCC 809
QY 178 GlyTyrGlyHis---GlyGluValSerSerSerLeuHisLysArgIle---HisProTyr 195
Db 810 CCGTACCATCATATGATGTCGCCAATCGTGAACCAATTGGCAAGCGGTCTCCATTTCCCTAT 869
QY 196 ProValSerGluProGlySerAlaLysTrpAspGluLysLys-----GluValSerTrp 213
Db 870 GTGAACCAATTGCGCAAAATCGTCAAGAGAGTTCTCTGTAGCATTTGGGAATGTTGCCCTGG 929
QY 214 LysGluArgMetAspAspTrpLysSerLysGln-----GlyIleLeu----- 227
Db 930 AAAGAGAGGTTGATGGCTGGGAAATGAAGCAGGACAAGGGGACGATTCCCATGACGAAT 989
QY 228 -----GlyGlyAlaAspProGluAspMetAspAla----- 238
Db 990 GGCACAAGCATTTGCTCCCTCTGAGGGTCGGGTGTGTGTGATATTGATGCACTCACTGAT 1049
QY 239 -----AspValAlaLeuAenAspGluAlaArgGlnProLeuSerArgLysVal 254
Db 1050 TACAACATGGAAGATGCGCTTATTGAACAGCAAGAACTCGACAGCCTCTATCTAGGAAGTT 1109
QY 255 SerIleAlaSerSerLysValAenProTyrArgMetValIleValArgLeuValVal 274
Db 1110 CCATCTCCTCTCCAGGATAAATCCATACAGATGGTCATTGTCTGCGATTGATTGT 1169
QY 275 LeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIleClyLeuTrp 294
Db 1170 CTAAGCATCTCTTGCACTACCGTATCAAAATCTGTGCGCAATGCAATCCCATTTATGG 1229
QY 295 LeuValSerIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPhePro 314
Db 1230 CTTCATCTGTATTATGTGAGATCTGTTGCTCTTTCTGTGATATTGATCATGTTCCCT 1289
QY 315 LysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArg 334
Db 1290 AAGTGGTTTCCAATCAACGGGAGAGCTACCTTTGATAGCTGGCATTAAGGTATGACCGG 1349
QY 335 GluGlyGluProSerLeuSerAlaValAspLeuPheValSerThrValAspProLeu 354
Db 1350 GAAGGTGAGCCATCTCAGTTGGCTGCTGTGATCTTTTCTGTCAGTACAGTCGACCAATG 1409

QY 355 LysGluProProLeuValThrAlaAenThrValLeuSerIleLeuAlaValAspTyrPro 374
Db 1410 AAGAGGCTCTCTTGTCTACTGCAATACCGTGTCTATCCATTCTTCTGTGGATTACCTT 1469
QY 375 ValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSer 394
Db 1470 GTGGATAAGGTCTCTTGTCTATCTGATGATGATGATGATGATGATGATGATGATGATG 1529
QY 395 LeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIle 414
Db 1530 CTAGCTGAGACTTCTGAGATTGCTAGAAATGGGTACCAATTGTTAAGAAGTACCAACTT 1589
QY 415 GluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysVal 434
Db 1590 GAACCTAGACTCTCTGAATGGTACTTCTCCAGAAAATTCATTACTTTGAAGGACAAAGTG 1649
QY 435 GluProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysVal 454
Db 1650 CACCTTTCATTTGTTAAAGACCGCGCGCCCATGAAGAGAGATATGAAGAATTCAAAGTT 1709
QY 455 ArgIleAenAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTyrIleMetLys 474
Db 1710 AGGTAAATGGCTTGTTCCTAAGGCACAGAAAGTCTCTGAGGAGGATGATCATGCAA 1769
QY 475 AspGlyThrProTrpProGlyAenAsnThrArgAspHisProGlyMetIleGlnValPhe 494
Db 1770 GATGGCACACCATGGCCAGGAAACAATACCMGGGACCATCTCGAATGATTCAGTTTTC 1829
QY 495 LeuGlyHisSerGlyGlyHisAspThrGluGlyAenGluLeuProArgLeuValTyrVal 514
Db 1830 CTTGGTCACAGTGGTGGCTTGATCTAGGCGCAATGAGCTACCCCGTTTGGTCTATGTT 1889
QY 515 SerArgGluLysArgProGlyPheGlnHisLysLysValAlaGlyAlaMetAsnAlaLeu 534
Db 1890 TCTCGTAAAGCGCTCTCGGATTCAGATCACAAAGAGCTGTGTCGATGAATGCTCTT 1949
QY 535 IleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAenLeuAspCysAspHis 554
Db 1950 GTTCGTCTCTCAGCTGTGCTTACCAATGGACAATACATGTTGATCTTGTATGATCAC 2009
QY 555 TyrIleAenAenSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnVal 574
Db 2010 TACATTAAACAACAGTAGGCTCTCAGGGAAGTATGCTTCTCTTATGGAACCTTAACCTA 2069
QY 575 GlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAsp 594
Db 2070 GGAAGGAGTGTCTCTACGTCAGTTTCCCAGAGATTCGATGGCATTCAGAGATGAT 2129
QY 595 ArgTyrAlaAenArgAenThrValPhePheAspIleAenMetLysGlyLeuAspGlyIle 614
Db 2130 CGATATGCCAACAGAACACCGCTGTTTTTCGATATTAACTTGAAGAGTCTTGTATGGCATC 2189
QY 615 GlnGlyProValTyrValGlyThrGlyCysValPheArgGlnAlaLeuTyrGlyTyr 634
Db 2190 CAAGGACAGTTTATGTCGGAACCTGGTGTGTTTTTCAACCGAACAGCTCTATATGGTTAT 2249
QY 635 AsnProProLysGlyProLysArgProLysMetValThrCysAspCysAspProCysPhe 654
Db 2250 GAGCCCCCAATTAAGCAGAGAGGGGTGTTCTTGTC-----TCATATGTGGC 2300
QY 655 GlyArgLysLys-----ArgLysHis 661
Db 2301 GGTAGGAAGAGGCAAGCAAAATCAAGAGGGCTCGGACAAGAGAGTTCGAGAAGCAT 2360
QY 662 AlaLysAspGlyLeuPro-----GluGlyThrAlaAspMet 673
Db 2361 GTGACAGCTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGGGAGTTGAAGCGCT 2420
QY 674 GlyValAspSerAspLysGluMetSerHisMetAenPheGluLysArgPheGly 693
Db 2421 CGATTGACGACGAGAAATCACTTCTTATGTCCTCAATGAGCCTGGAGAGAGATTGGC 2480
QY 694 GlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSer 713

Job time : 1012 secs

2481	Db		CAGTCCG	CAGGTTGT	GTGGCT	CCAC	TCTGAT	GAGTAT	GGTGGT	GTCTCT	CAGTCCG	CA	2540
714	Qy		SerPro	AlaLeu	LeuLeu	LeuLeu	LeuLeu	LeuLeu	LeuLeu	LeuLeu	LeuLeu	LeuLeu	733
2541	Db		ACTCG	GAGTCT	CTCT	CTT	GAAAG	AGACT	TATCC	ATGTTAT	AAGCTGT	GGCTAT	2600
734	Qy		ThrAsp	TrpGly	LeuGlu	LeuLeu	GlyTyr	TrpPhe	GlySer	LeuThr	GluAsp	LeuLeu	753
2601	Db		ACTGAA	TGGGAA	CTGAG	ATCGG	TGGAT	CTTAC	CGTTC	TGTG	CAGAGA	CAATTC	2660
754	Qy		GlyPhe	LeuMet	HisCys	ArgGly	TrpArg	SerVal	TyrCys	MetPro	LeuLeu	AlaAa	773
2661	Db		GGATT	CAAGAT	GCAC	CGCG	GAGGCT	GGCTGC	ATCTAC	TGATG	CCCAAG	CGCCAG	2720
774	Qy		PheLeu	GlySer	AlaPro	IleAsn	LeuSer	AspArg	LeuAsn	GlnVal	LeuArg	TrpAla	793
2721	Db		TTCAAG	GGGTCT	GC	CCCC	CAATCT	TTTC	GACG	CGCTG	TAAC	CAAG	2780
794	Qy		LeuGly	SerVal	GluIle	PhePhe	SerArg	HisSer	ProLeu	LeuTyr	GlyTyr	LeuAsn	813
2781	Db		CTTGG	GTCT	CCGTGG	AGATCT	CTT	CAGC	CGCACT	TGCC	CCCTGT	GTGAC	2837
814	Qy		GlyAsn	LeuLeu	ValTrp	LeuGlu	ArgPhe	AlaTyr	IleAsn	ThrThr	IleTyr	ProPhe	833
2838	Db		GGCGCG	CTCA	GTTC	TCTGG	AGAGAT	TCCG	GTATCA	TCAAC	CAAC	CTATCA	2897
834	Qy		SerLeu	ProLeu	LeuAla	TyrCys	ThrLeu	ProAla	ValCys	LeuLeu	ThrGly	LeuPhe	853
2898	Db		TCCAT	CCGCTT	CTCT	CACT	TACTG	ATCT	TGCC	CGCCAT	CTGT	TGCT	2957
854	Qy		IleMet	ProSer	IleSer	ThrPhe	AlaSer	LeuPhe	PheIle	AlaLeu	PheMet	SerIle	873
2958	Db		ATCAT	CCAGAT	CTAC	GAATCT	TCCG	CAGAT	CTTGG	TTCATCT	CCCTCT	TTCAT	3017
874	Qy		PheAla	ThrGly	IleLeu	GluMet	ArgTrp	SerGly	ValSer	IleGlu	LeuTrp	TrpArg	893
3018	Db		TTGC	CAAC	GGGCAT	CTCT	GGAGAT	TAGT	GTGAG	CGGGT	TGGCAT	TCA	3077
894	Qy		AsnGlu	GlnPhe	TrpVal	IleGly	GlyVal	SerAla	HisLeu	PheAla	ValVal	GlnGly	913
3078	Db		AACG	AGCAGT	CTCT	GGTGT	GATCG	GGGCAT	CTCT	CGCG	CACTCT	TTC	3137
914	Qy		LeuLeu	LeuVal	LeuAla	GlyIle	AspThr	AsnPhe	ThrVal	ThrSer	IleAla	ThrGly	933
3138	Db		CTGCT	CAAGT	GTGT	TGGCG	CGCAT	CGA	CAAC	AACTT	TCCCGT	CACT	3194
934	Qy		AspGlu	AspAsp	GluPhe	AlaGlu	LeuTyr	AlaPhe	IleVal	SerTrp	ThrLeu	LeuIle	953
3195	Db		GACG	AGAC	CGGCAT	CTCG	GGAGCT	GTAT	GTTC	AA	GTGG	AGAC	3254
954	Qy		ProThr	ThrLeu	LeuLeu	IleIle	AsnVal	IleGly	ValVal	AlaGly	IleSer	AspAla	973
3255	Db		CCCAC	CACT	CTCT	GATCAT	CAAC	CTGGT	CGCGT	CTCG	CGCG	CACT	3314
974	Qy		AsnAsn	GlyTyr	GlnSer	TrpGly	ProLeu	PheGly	IleLeu	PhePhe	AlaPhe	TrpVal	993
3315	Db		AAACG	CGGAT	TACAGT	CTGG	GGCG	CGCTCT	TTCG	GAAGCT	CTTCT	TTCG	3374
994	Qy		IleVal	HisLeu	TyrPro	PheLeu	LeuVal	GlyLeu	MetGly	ValArg	GlnAsn	ArgThr	1013
3375	Db		ATGCT	CACT	GTAT	CCCGT	TCTCA	AGGG	CGCTCAT	GGCG	AGGAC	CGAC	3434
1014	Qy		ValVal	ValIle	TrpSer	IleLeu	LeuAla	SerIle	PheSer	LeuLeu	TrpVal	ArgIle	1033
3435	Db		ATCGT	CGTCT	GTCT	GGCCAT	CTCT	GTCT	GGCGT	CCATCT	TCTCT	TCT	3494
1034	Qy		AspPro	PheIle	ValArg	ThrGly	SerPro	AspVal	ArgGln	CysGly	IleAsn	Cys	1052
3495	Db		GACCC	TTAC	CAAC	CCCG	CGTCACT	TGGC	CGCGAT	TACCA	CGAC	GTGG	3551

This Page Blank (uspto)


```
Qy 669 -----GlyThrAlaAspMetGlyValAspSerAspIysGlu 680
Db 2232 CQTGATATGCTCTTTGGTGAATTCACGAAGCTGCTCCAGGAGCTGAGAAATGAAAAGGCC 2291
Qy 681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
Db 2292 GGTATTGTAAATCAACAAAATAGAAAGAAATTTGGCCAAATCTCTGTTTTTTTGTACA 2351
Qy 701 SerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeuLys 720
Db 2352 TCCACACTTCTCGAGAAATGTTGAACCTTTGAAGAGTGCAGGCTCTGCTTCTTTTGAA 2411
Qy 721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTyrPheGlyLeuGluLeu 740
Db 2412 GAAGCTATACATGTCATTAGTTGTGTTATGAAGACAACAGACACTGGGGAAGAGATT 2471
Qy 741 GlyTyrIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
Db 2472 GCCTGGATCTATGGATCAGTTACAGAGATATCTAACTGGTTTCAAGATGCAATGTGCAT 2531
Qy 761 GlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
Db 2532 GGTTCGGGTCAATTTACTGCATACCTTAACGGGTTGCATTCAAGAGTTCTGCACCTCG 2591
Qy 781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTyrPheAlaLeuGlySerValGluIlePhe 800
Db 2592 AATCTTTTCAGATCGTCTTACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC 2651
Qy 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
Db 2652 TTCACAAATCATGTCCTCTTTGGTATGGGTAT---GGTGGCGGTCTGAAATTTTGGAA 2708
Qy 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr 840
Db 2709 AGATTTTCTACATCAATCCATCGTATCTCTGGACATCTATTCCCTCTTGGCTTAC 2768
Qy 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTACATTTGCCCTGCCATCTGTTTATGACAGGAAATTTATCACTCCAGAGCTGAATAAT 2828
Qy 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 2829 GTTGCACGCTGTGGTGTATGTCATCTTTTATCTGCATTTTGTCTACAGCATCTCAGAA 2888
Qy 881 MetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIle 900
Db 2889 ATGAGATGGAGTGGTGTGGATTCATGCTGCTGGAGGAATGACAGATTTCTGGGTCAAT 2948
Qy 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
Db 2949 GGAGGTGTCTCTCACACTCTTGTCTGTCTCCAGGACATCTCTCAAGGTCTATAGCTGGT 3008
Qy 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAspGluPheAla 940
Db 3009 GTTGATCAAGCTTCACCGTGACATCAAG-----GGTGGAGATGATGAGGAGTTCTCA 3062
Qy 941 GluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleIle 960
Db 3063 GAGCTATATACATTCAAATGACTACCTTATTGATACCTCTTACACCTTGTCTTCTATTG 3122
Qy 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
Db 3123 AACTTCATTGTGTGCTGGCTGCTTCAATATGCGATCAATAACGGATATATGAGTCATGG 3182
Qy 981 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1000
Db 3183 GCCTCCCTCTTTGGGAAGCTATCTTTGCAATTTGGGTGATGTGTCATCTTATCCCTTT 3242
Qy 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
Db 3243 CTCAAAGGTTTGGTTGGAGGCGCAAAACAGCACCAACAGATGTGTCATCTGTCTGGTCAAT 3302
```

```
Qy 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 3303 CTGTGGCTTCAATCTTCTGCTCTCTTGGGTGGATGTGATCTCTTCTTCCGGAAGGAT 3362
Qy 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3363 GATGTCCTGCTTCTTGAGGAGTGIGGTTGGATTGC 3398

RESULT 2
US-10-160-719A-37
; Sequence 37, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)....(3401)
US-10-160-719A-37

Alignment Scores:
Pred. No.: 0 Length: 3969
Score: 4073.50 Matches: 757
Percent Similarity: 79.68% Conservative: 129
Best Local Similarity: 68.08% Mismatches: 139
Query Match: 72.12% Indels: 87
DB: 4 Gaps: 15

US-10-627-132-30 (1-1052) x US-10-160-719A-37 (1-3969)
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 144 ATGGAGCGAGCGCGCGGCTGTGTGGCGGCTCCCAACACCGCAACGAGCTGTGTCTATC 203
Qy 21 Arg-----GlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGlu 38
Db 204 CGCCGCGAGCGGATCCCGGCCCGAGCCCGCGGAGCAGACGCGGCGGTGTCCAG 263
Qy 39 IleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGlu 58
Db 264 ATTTGCGGACGACGCTCGGCTTTCGCCCGCGGCGGACCCCTTCGTGGCGTCAACGAG 323
Qy 59 CysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsn 78
Db 324 TGCCTCTTCCCGCTCTCGCGGACTGTCTACAAATACGAGCGCGGAGGCGCACGCAAC 383
Qy 79 CysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAsp 98
Db 384 TGCCCCCAGTGCAAGACTCGATACAGGCTCAAGGGCTCCCAACGCTGTGACCGGTGAC 443
Qy 99 AspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 118
Db 444 GAGGAGGAGGAGCGCGTCTGATGACCTGACAAACGAGTTCAACTGGGACGCGCATGCTCG 503
Qy 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
Db 119 GlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
```

504 CAG-----TCTGTGGCCGAGTCCATGCTCTACGGC 533
Db
139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
Qy
534 CACATGAGCTACGCCCGTGA-----GGTGACCTTAATGGCGGCCACACGCTTTC 584
Db
159 ProfileIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGly 178
Qy
585 CAGCTC-----AACCCCAATGTTCCACTC-----CTCACCAACGGG 620
Db
179 TyrGlyHisGlyGluValSerSerSerLeuHis----- 189
Qy
621 CAATGTGTGATGATACATCCACCGGAGCAGCAGCGTGTGCTTCTTTCATGGGTGT 680
Db
190 -----LysArgIleHisProTyrProValSerGluProGly-----SerAla 203
Qy
681 GGGGGAAGAGGATACATCCCTTCTTATGCGGATCCACGCTTACCTGTGCAACCCAGG 740
Db
204 LysTrpAspGluLysGlu-----ValSerTrpLysGlu 215
Qy
741 TCTATGACCCATCCAGGATCTTGTGCATATGGGTATGTTAGTGTGTGGAAGAA 800
Db
216 ArgMetAspAspTrpLysSerLysGln-----GlyIleLeuGlyGly 229
Qy
801 CGGATGAGAAATTGGAAGCAGAGCAAGAGAGGATGCACACGCGCGGAATGATGTGTGT 860
Db
230 GlyAlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaAspGlnPro 249
Qy
861 GGTGATGATGTGAC-----GATGCTGATCTACCACTTAATGGATGAAGCAAGCAACA 914
Db
250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
Qy
915 CTGTCCAGGAATAATCCATTCCTCAAGCAGCAATTAATCCATATAGATGATATCAAT 974
Db
270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
Qy
975 ATTCGGCTTGTGTTTGGGGTCTCTTCCACTACCGAGTGCATCCGGTGAATGAT 1034
Db
290 AlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIle 309
Qy
1035 GCATTTGCTTGTGGCTCATATCTGTTATCTGTGAATCTGGTTTGGCATGTCTTGGATT 1094
Db
310 LeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSer 329
Qy
1095 CTTGATCAATCCCAAGTGTTCCTATTCAGAGAGACTTACCTAGACCGCTGTCA 1154
Db
330 LeuArgTyrGluArgGluGlyProSerLeuLeuSerAlaValAspLeuPheValSer 349
Qy
1155 CTGAGTTCCGACAAGGAGCCGCACTCTCAACTTGTCTCAATGTATTTCTTGTCAGT 1214
Db
350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu 369
Qy
1215 ACGGTGTATCCCTTAAAGGAACCTCCTTGTGTCAACAACAAATCTGTTCTATCTATCCTT 1274
Db
370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 389
Qy
1275 TCGGTGAATATCCCTGTTGATAGGTTTCTTGTCTATGTTTCTGATGTTGTGCAATG 1334
Db
390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCys 409
Qy
1335 CTAACGTTTGAAGCATATCTGAAACATCTGAATTTGCAAGAAATGGGTTCCITTTCTGC 1394
Db
410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
Qy
1395 AAACCGGTACAATATTGAACCTCGCGCTCCAGGTGGTACTTCCAAACAGAAATAGACTAC 1454
Db
430 LeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyr 449
Qy
1455 TTGAAGACACAGTGGCGCAACACTTGTAGGGAGAGGAGCAATGAAGAGAGATAT 1514
Db
450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
Qy
1515 GAGGAATCAAGGTGAGAATCAATGCCTTAGTTTGGCAAAAGGCCCAAGAAAGTTCTCGAAGA 1574
Db

470 GlyTrpIleMetLysAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 489
Qy
1575 GGNATGACATGCAAGATGGNACCCCTGGCTGGAAACAATGTTCTGTGATCATCTCTGGA 1634
Db
490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
Qy
1635 ATGATTCAAGTCTCTCTTGGCCAAAGCGGAGGCTTGTGACTGTGAGGAAATGAATGACCA 1694
Db
510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysValAlaGly 529
Qy
1695 CGATTGGTTTATGTTTCTAGAGAAACGACCAAGGCTATAACCATCATATAAGAAAGCTGT 1754
Db
530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
Qy
1755 GCTATGATGATGCTGTGAGTCTCTGCTGTACTAAACAATGCTCCATATTTGTTAAAC 1814
Db
550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
Qy
1815 TTGGATTGTGATCACTACATCAACAACAGGCTATAAAGGAAGCAATGTGTTTATG 1874
Db
570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
Qy
1875 ATGACCCCTTACTAGGAAAGAGGTTTGTATGTACAGTCTCCTCAAGATTTGATGGG 1934
Db
590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
Qy
1935 ATTGATCGCATGACCGATATGTAACCGGAATGTTGTCTTTTGTATATCAACATGAA 1994
Db
610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGln 629
Qy
1995 GGTTCGATGGTATTTCAGGGTCCAAATTTATGTTGGTACTCGATGTGTATTTAGAACGAG 2054
Db
630 AlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAsp 649
Qy
2055 GCATTATATGTTATGATGCCCCAAA---ACAAGAAAGCCCAATCAAGAGCTTGCAC 2111
Db
650 -----CysCysProCysPheGlyArgLysLysArgLysHisAla 662
Qy
2112 TGCTGGCCCAAGTGGTCTTTGCTGTGCTGCTTTGGCAATAGGAAGCAAAAGAGACT 2171
Db
663 LysAspGlyLeuProGlu----- 668
Qy
2172 ACCAAACCCAAACAGAGAAAGAAAGTATTATTTTCAAGAAAGAGAACCAATCC 2231
Db
669 -----GlyThrAlaAspMetGlyValAspSerAspLysGlu 680
Qy
2232 CCTGCATATGCTCTTGTGTAATTTGCAAGAGCTGCTCCAGAGCTGAGAATGAAGAGCC 2291
Db
681 MetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThr 700
Qy
2292 GGTATGTAAATCAACAAAAATTAGAAAGAAATTTGCCCAATCTCTGTTTTTTGTATCA 2351
Db
701 SerThrLeuMetGluGlyGlyValProProSerSerSerProAlaAlaLeuLeuLys 720
Qy
2352 TCCACACTTCTCGAGAATGTTGGAAACCTTCAAGAGTGCAGTCTCTCTCTTTGAAA 2411
Db
721 GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu 740
Qy
2412 GAAGCTATACATGTCATGTGTTGGTTATGAAGACAGACAGACTGGGGAAGAAAGAT 2471
Db
741 GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg 760
Qy
2472 GGCTGGATCTATGATCAGTTACAGAGATATTTCTTAAGTGGTTTCAAGATGCAATGTCT 2531
Db
761 GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle 780
Qy
2532 GGTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTTCAAAAGGTTCTGCACCTCTG 2591
Db
781 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhe 800
Qy
2592 AATCTTTCAGATCGTCTTCACCAGGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC 2651
Db

```

QY 801 PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu 820
Db 2652 TTCAGCAATCATGGCCCTTTGGTATGGGTAT---GGTGGCGGTCTGAATAATTTTGGAA 2708
QY 821 ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuSerLeuLeuAlaTyr 840
Db 2709 AGATTTCTACATCAATCCATCCATCGTGTATCTCTGGACATCTATCCCTCTTGCTTAC 2768
QY 841 CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr 860
Db 2769 TGTACATTCCTGCTGCCATCTGTTTATTCAGCGGAAATTTATCACTCCAGAGCTGAATAAT 2828
QY 861 PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 2829 GTTGCCAGCTGTGGTGTATGTCATCTTTATCTGATTTTGTACGAGCATCTCTAGAA 2888
QY 881 MetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluGluPheTrpValIle 900
Db 2889 ATGAGATGGAGTGGTGTGGATTCATGACTGCTGGAGGAATGACAGATTCTGGGTCAIT 2948
QY 901 GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly 920
Db 2949 GGAGGTGTCTCTCACACCTCTTGTCTGTGTCCAGGACTTCTCAAGTCTATGCTGT 3008
QY 921 IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla 940
Db 3009 GTTGATACAAGTTCACCGTGCATCAAAAG-----GGTGGAGATGATGAGGAGTCTCA 3062
QY 941 GluLeuTyrAlaPheLeuTyrTrpThrLeuLeuIleProProThrThrLeuLeuIleIle 960
Db 3063 GAGCTATATACATTCAAATGGACTACCTTATGATACCTCTCCACCTTGCTCTATGT 3122
QY 961 AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp 980
Db 3123 AACTTCATTGGTGTGCTGCTGGCTTCAATGCGATCAATAACGATATGAGTCAATG 3182
QY 981 GlyProLeuPheGlyLysLeuPheAlaPheTrpValIleHisLeuTyrProPhe 1000
Db 3183 GGCCCCCTCTTTGGGAAGCTATCTTGTGATTTGGGTGATGTCATCTTATCCCTTT 3242
QY 1001 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle 1020
Db 3243 CTCANAGTGTGGTGGAGGAAACAGGACACCAAGATTCATGCTCTGGTCCAT 3302
QY 1021 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr 1040
Db 3303 CTGCTGGCTTCAATCTCTGCTCTCTTGGGTGGGATTGATCTCTTCTTCTCGAAGGAT 3362
QY 1041 LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3363 GATGTCCTGCTCTTTGAGGAGTGTGGTTGGATTGC 3398

```

RESULT 3

```

US-10-160-719A-1
; Sequence 1, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1

```

```

; LENGTH: 3568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)...(3239)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3487)...(3487)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(3568)
; OTHER INFORMATION: n = A,T,C or G
US-10-160-719A-1

```

Alignment Scores:

```

Pred. No.: 0 Length: 3568
Score: 4005.00 Matches: 743
Percent Similarity: 79.04% Conservative: 128
Best Local Similarity: 67.42% Mismatches: 141
Query Match: 70.91% Indels: 90
DB: 4 Gaps: 16

```

US-10-627-132-30 (1-1052) x US-10-160-719A-1 (1-3568)

```

QY 14 ArgAsnGluLeuValLeuIleArg-----GlyHisGluAspProLysProLeuArgAla 31
Db 12 CGTCGGAGCTCGTCTGTCATCGCGCGATGCGGAGCCAGGCGCGAAGCCCATGGACCAG 71
QY 32 LeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAsp 51
Db 72 CGGAACGCCAGGTGTGCCAGATTTGGCGGACGACGCTGGGGCGCAACCCCGACGGGAG 131
QY 52 LeuPheValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGlu 71
Db 132 CCTTGTGGCTGCAAGAGTGGCCCTTCCCATCTCCCGGACTGTCTAGGATACGAG 191
QY 72 ArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGly 91
Db 192 CGCCGCGAGGCGACGCGAAGTCCCGCCAGTGCAGACCCGCTTCAAGCGCTTCAAGGG 251
QY 92 SerProArgValAlaGlyAspAspGluAspIleAspAspLeuGluHisGluPhe 111
Db 252 TGCOCGCGCGTGGCGGGGACGAGGAGGAGGAGCGCTCGACGACCTGGAGAACGAGTTC 311
QY 112 AsnIleAspAspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIle 131
Db 312 AACTGGAGCGACAGCAGCTCCAGTACCTC----- 344
QY 132 ThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAsp---AspGlyAsp 150
Db 345 GCGGAGTCCATGCTCCAGCCCATGAGCTACGCGCGCGCGCGGACCTCGACGGCGTG 404
QY 151 GlyAsnAsnThrProGlnIleProProIleIleThrGlySerArgSerValProValSer 170
Db 405 CCGCAGGCGATTCCACCCCATCTCCC-----AATGTTCCCTC--- 440
QY 171 GlyGluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeuHis--- 189
Db 441 -----CTCACCAACGACAGATGCTGATGATCCCGCGGACGAGCACGCC 488
QY 190 -----LysArgIleHisProTyrProValSerGlu 199
Db 489 CTTGTGCCCTCGTTCGTGGTGGGGGGGAGAGGATTACCCCTCTCCCTACCGGAT 548
QY 200 ProGly-----SerAlaLysTrpAspGluLysGlu----- 210
Db 549 CCCAACCTTCTCTGTGTCAACCGAGGTCTATGACCCCTTCCAGGATCTCGCGCATATGC 608
QY 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLysGlnIleLeu 227
Db 609 TACGGGAGCGGTAGTAGCAAGGAGGATGGAGAGTGGAGAGTGGAAAGACGAGGAGGATG 668

```

QY 228 -----GlyGlyValaAspProGluAspMetAspAlaAspVala 241
DB 669 CACCAGACGAGAACGATGGCGGCGCATGGTGTAT-----GATGCAGATCTACCA 722
QY 242 LeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLysVal 261
DB 723 CTAATGGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCCTTCAAGCCAAATC 782
QY 262 AsnProTyrArgMetValIleValValArgGlnValValLeuAlaPhePheLeuArgTyr 281
DB 783 AACCCCTATAGATGATTAATAATTCGGCTAGTGGTTTGTGTTCTTCTCCACTAC 842
QY 282 ArgIleLeuHisProValProAspAlaIleGlyLeuTyrLeuValSerIleIleCysGlu 301
DB 843 CGAGTGATGATCCCGTGCCTGATGCAATTCCTTATGGCTCATATCTGTGATCTGTGAA 902
QY 302 IleTyrPheAlaIleSerTyrIleLeuAspGlnPheProLysTyrPheProIleAspArg 321
DB 903 ATTTGGTTTGGCATGTCTTGGATCTTGACCAAGTTTCCAAAGTGGTTTCTTATCGAGAG 962
QY 322 GluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeu 341
DB 963 GAAACCTATCTGACCGGCTGAGTTAAGGTTTGACCAAGGAAGGCGATCTTCTCAACTC 1022
QY 342 SerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuValThr 361
DB 1023 GCCCCTGTGTATTTCTTGTGATGATCGGTTGATCCCTTGAAGGAACCTCCATTTGGTCACT 1082
QY 362 AlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyr 381
DB 1083 GCTAATACTGTCTATCTATCTTCTCGGTGATTCATCCAGTTTGATGAAGGTTTTCATGCTAC 1142
QY 382 ValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGluPhe 401
DB 1143 GTTTCGTGATGATGGTGTGCCATCTGACATTTGAAGCATTTGTCTGAACATCTGAATTT 1202
QY 402 AlaArgLysTyrValProPheCysLysLysPheGlyIleGluProArgAlaProGluPhe 421
DB 1203 GCAAGAAATGGGTTCCCTTCTGCAAAAGATATAGCCTTGAGCCTCGTGTCTCCAGAGTG 1262
QY 422 TyrPheSerLeuLysValAspTyrIleLysLysAspLysValGlnProThrPheValGlnGlu 441
DB 1263 TACTTCCAAACAGATAGACTACCTGAAAGACAAAGGTGGCGCCAAACTTTGTTAGAGAA 1322
QY 442 ArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuValAla 461
DB 1323 CGGAGCAATGAAGAGAGATGATGAGGAATTCAGGTCAGAAATCAATGCCCTCTGTGTCT 1382
QY 462 LysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThrProTyrProGly 481
DB 1383 AAAGCCCAAAAGGTTCCCTGAGGAGGATGGACAATGCAGGATGGAACCTCCATGCCCGGA 1442
QY 482 AsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHis 501
DB 1443 AATAATGTCGCGTATCATCTGGAATGATTCAGGTTTTCCTTGGTCAAAAGTGGTGGCCAT 1502
QY 502 AspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGly 521
DB 1503 GATGGGAAGGAAATGAGCTGCTCGATGGTGTATGTTTGAAGTGTTCGTTTCAAGAGAAAACGGCCAGGC 1562
QY 522 PheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeu 541
DB 1563 TACAACCATCACAAGAGGCTGGTCTATGAATGATGATGGTCCGAGTCTCTGCTGTACTA 1622
QY 542 ThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAla 561
DB 1623 ACTAATGCTCTCTATTGTGAACCTTGGATTTGGATTCATATATCAATTAATAGTAGGCT 1682
QY 562 IleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyrVal 581
DB 1683 ATAAAGGAAGCAATGTGTTTATGATGGATCTCTTGTGGAAAGAAAGTTTGTCTATGTG 1742
QY 582 GlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThr 601

DB 1743 CAGTTTCTCAAAAGATTGATGGGATTGATCGCATATGATCGTATACAGAAATGTT 1802
QY 602 ValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGly 621
DB 1803 GTCTTTTTCGATATCAACATGAAGGTTTGGATGGTATCAGGCGCCCAATTAATGTGGGT 1862
QY 622 ThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLys 641
DB 1863 ACTGGATGTCTTCAGNAGGCGAGGCATATATGGCTAGCATGCTCCCAA---ACAAAG 1919
QY 642 ArgProLysMetValThrCysAsp-----CysCysProCysPhe 654
DB 1920 AAGCCACCATCAAGAACTTGCACACTGCTGGCCAAAGTGGTGCATTTGCTGTGCTGTTT 1979
QY 655 Gly---ArgLysLysArgLysHisAlaLysAspGlyLeuProGlu----- 668
DB 1980 GGTAACAGGAAGAACCAAGAAAGAACCAAGACCTCTAAACCTAAATTTGAGAAGATAAAG 2039
QY 669 -----GlyThr 670
DB 2040 AAACCTTTTAAAGAAAAGGAAATCAAGCCCTGCATATGCTCTTGTGTGAAATTTGATGAA 2099
QY 671 AlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLys 690
DB 2100 GCCGCTCCAGGAGCTGAAATGAAGGCTAGTATTGTAAATCAACACAGAAAGTTGGAAAG 2159
QY 691 ArgPheGlnSerAlaPheValThrSerThrLeuMetGluGluGlyGlyValPro 710
DB 2160 AAATTTGGCCAGCTTCTCAGTTTGTGTCATCCACACTTCTTGAGAAATGGTGGNACCTG 2219
QY 711 ProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 730
DB 2220 AAGAGTGCCAGTCCAGCTTCTTCTGAAGGAAGCTATACATGTCATCAGTTGTGGATAT 2279
QY 731 GluAspLysThrAspTyrGlyLeuGluLeuGlyTyrIleTyrGlySerIleThrGluAsp 750
DB 2280 GAAGACAAAACAGCTGGGGAAGATATTTGGTTGGATTTATGGATCAGTCACAGAAAGAT 2339
QY 751 IleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLys 770
DB 2340 ATTTCTACTGGTTTAAGATGCACTGCCATGGTTGGCGGTCAATTTACTGCATACCTAAA 2399
QY 771 ArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 790
DB 2400 CGGCGCCCTTCAAGGTTCCGACACTCTCAATCTTTCCGATCGTTTTTCCACAGGTTCTT 2459
QY 791 ArgTyrAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGly 810
DB 2460 CGGTGGGCTCTTGGTTCAATTTGAAATTTTGTTCAGCAACCACTGCCCTCTCTGGTATGGG 2519
QY 811 TyrLysAsnGlyAsnLeuLysTyrLeuGluArgPheAlaTyrIleAsnThrIleTyr 830
DB 2520 TAT---GGTGGTGACTAAAGTTTCTGGAAAGGTTTTTCGTACATTAATCTCATCGTATAC 2576
QY 831 ProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThr 850
DB 2577 CTTGGACATCTATCCCGCTCTTGGCCTATTGCACTATGCTGCCATCTGCTGCTGACA 2636
QY 851 GlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPhe 870
DB 2637 GGGAAATTTATCAGCCAGCAGCTTAACAATGTTGCCAGCCTCTCGGTTTCATGCTCACTTTC 2696
QY 871 MetSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSerIleGluGlu 890
DB 2697 ATCTGCATTTTGTCTAGCAGCATCTCGAAATAGATGGAGTGTGTAGGCATCGATGAC 2756
QY 891 TrrTyrArgAsnGluGlnPheTrrValIleGlyValSerAlaHisLeuPheAlaVal 910
DB 2757 TGGTGGAGAAACGAGCAGTTTTTGGTCACTTGGAGGCGGTCTTTCACATCTCTTGTGCTGTG 2816
QY 911 ValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 930

```
Db      2817 TTCCAGGGACTCCTCAAGGTCTAGCTGGTGTAGACACGAGCTTCACTGTGACATCCAAG 2876
Qy      931 AlaThrGlyAspGluAspGluPheAlaGluLeuTyxAlaPheLysTrpThrThrLeu 950
Db      2877 -----GGCGGAGACGACGAGGAGTTCTCAGAGCTGTACACATTCAATGGACGACCCCTT 2930
Qy      951 LeuLeuProProThrThrThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 970
Db      2931 CTGATACCTCCGACAAACCTGCTCTACTCACTTCAATGGAGTGGTAGCTGACATCTCC 2990
Qy      971 AspAlaIleAsnAsnGlyTyxGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAla 990
Db      2991 AATGCGATCAACACGATATGATCATATGGGCCCCCTGTTGCGGAGCTCTCTTTGCA 3050
Qy      991 PheTrpValIleValHisLeuTyxProPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1010
Db      3051 TTTTGGGTGATCGTCCATCTTTACCGTTCCTCAAGGGTCTGTTGGAGGCGACAGCAG 3110
Qy      1011 ThrProThrValValValIleTrpSerIleLeuLeuLeuLeuLeuLeuLeuLeuLeuTrp 1030
Db      3111 AGCCCAACGATTGTCTATGTGTGTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3170
Qy      1031 ValArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIle 1050
Db      3171 GTCCCGATCGACCCGTTCTTGGAGAGATGATGTGTCCTGTTGGAGAGTGTGTCTG 3230
Qy      1051 AsnCys 1052
Db      3231 GATTGC 3236
```

RESULT 4

```
US-10-160-719A-25
; Sequence 25, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Drugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3813
```

```
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)...(3496)
US-10-160-719A-25
```

```
Alignment Scores:
Pred. No.: 0 Length: 3813
Score: 3974.50 Matches: 752
Percent Similarity: 78.53% Conservative: 122
Best Local Similarity: 67.57% Mismatches: 158
Query Match: 70.37% Indels: 81
DB: 4 Gaps: 17
```

```
US-10-627-132-30 (1-1052) x US-10-160-719A-25 (1-3813)
Qy      1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGlnLeuValLeuIle 20
Db      215 ATGGAGGCTAGCGCGGGGCTGGTGGCCGGCTCGCATNACCGAACGAGCTGGTGGTGATC 274
```

```
Qy      21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnVal----- 36
Db      275 CGC-----CGCGACCGCGAGTCGGGAGCCGCGCGCGCGCGCGCGCGCGCGCGAG 328
Qy      37 -----CyeGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheVal 54
Db      329 GCGCGGTGCGCATATATGCGCGACGAGTTCGGGGTGGGGTTCGACCGGGAGCCCTTCGTG 388
Qy      55 AlaCysAsnGluCysGlyPheProValCysArgProCysTyxGlnTyxGluArgArgGlu 74
Db      389 GGTGTCAACGAGTGGCGCTTCCCGCTGCGCGCGCTGCTACAGTACGAGCGCGCGAG 448
Qy      75 GlyThrGlnAsnCysProGlnCysLysThrArgTyxLysArgLeuLysGlySerProArg 94
Db      449 GCGTCGCAAGCGTGGCGCGAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 508
Qy      95 ValAlaGlyAspAspAspGluLeuAspIleAspLeuGluHisGluPheAsnIleAsp 114
Db      509 GTGGCGCGCGAGGAGGAGGAGCGCGCTGCGAGACCTGGAGGGCGAGTTCGGCGCTGCGAG 568
Qy      115 AspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAla 134
Db      569 GACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTCC 613
Qy      135 MetLeuHisGlyArgMetSerTyxGlyArgGlyProAspAspGlyAspGly---AsnAsn 153
Db      614 ATGCTCAGCGCGCAGATGAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 673
Qy      154 ThrProGlnIleProProIleIleThrGlySerArgSerValProValSerGlyGlu--- 172
Db      674 GTCCCCAACGTGCGCGCTCCTACCAACGCGCGAGATGGTTGATGACATCCCGCGCGAGCAG 733
Qy      173 -----PhePro-----IleThrAsnGlyTyxGlyHisGlyGluValSerSer 187
Db      734 CACGCGCTCGTCCCTCTACATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 778
Qy      188 LeuHisLysArgIleHisProTyxProValSerGluProGly-----SerAla 203
Db      779 -----AAGAGGATCCACCGCTCTTCGAGATCCCAACCTTCCAGTGCACCGCAGA 832
Qy      204 LysTrpAspGluLysLysGlu-----ValSerTrpLysGlu 215
Db      833 TCCATGGACCGCTCCAAGGATCTGCGCGCTACGGATATGGCAGCGTGGCGCTGGAGGAG 892
Qy      216 ArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGly 230
Db      893 AGAATGGAGGGCTGGAAGCAGAGCAGGCGCGCTGCGACATGTCCAGGACGAGGTTGCG 952
Qy      231 AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeu 250
Db      953 GGTGATTGGGATGGCGAGCATGTCCTCACTAATGGATGAAGCTAGGCGAGCATTTG 1012
Qy      251 SerArgLysValSerIleAlaSerSerLysValAsnProTyxArgMetValIleVal 270
Db      1013 TCCAGAAAAGTCCCTATATCATCAAGCGAATTAATCCCTACAGATGATATCGTTATC 1072
Qy      271 ArgLeuValValLeuAlaPhePheLeuArgTyxArgIleLeuHisProValProAspAla 290
Db      1073 CGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1132
Qy      291 IleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpLeu 310
Db      1133 TTTGCAATTTGGCTCATATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCT 1192
Qy      311 AspGlnPheProLysTrpPheProIleAspArgGluThrTyxLeuAspArgLeuSerLeu 330
Db      1193 GATCAGTTCCCAAGAGTGGCTTCCATCGAGAGAGAGACTTACCTGGAGCGGTTTGCATCA 1252
Qy      331 ArgTyxGluArgGluGlyProSerLeuLeuSerAlaValAspLeuPheValSerThr 350
Db      1253 AGTTTTCAGAGAGAGAGTCAACCTCTCAGCTTGTCTCCATCGACTCTTCTTGTCTAGTAGC 1312
Qy      351 ValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAla 370
```


RESULT 5

US-10-160-719A-45
 ; Sequence 45, Application US/10160719A
 ; Patent No. 6803498
 ; GENERAL INFORMATION:
 ; APPLICANT: Dhugga, Kanwarpal S.
 ; APPLICANT: Helentjaris, Timothy G.
 ; APPLICANT: Bowen, Benjamin A.
 ; APPLICANT: Wang, Xun
 ; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
 ; FILE REFERENCE: 0864C
 ; CURRENT APPLICATION NUMBER: US/10/160,719A
 ; PRIORITY FILING DATE: 2002-06-03
 ; PRIOR APPLICATION NUMBER: US 60/096,822
 ; PRIORITY FILING DATE: 1998-08-17
 ; PRIOR APPLICATION NUMBER: US 09/371,383
 ; PRIORITY FILING DATE: 1999-08-06
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 45
 ; LENGTH: 3813
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (215)...(3496)
 ; US-10-160-719A-45

Alignment Scores:

Pred. No.: 0 Length: 3813
 Score: 3974.50 Matches: 752
 Percent Similarity: 78.53% Conservative: 122
 Best Local Similarity: 67.57% Mismatches: 158
 Query Match: 70.37% Indels: 81
 DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-45 (1-3813)

Qy	1	MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuile	20
Db	215	ATGAGGCTAGCGCGGGCTGGTGGCGCTGCATACCGGACGAGCTGGTGTGATC	274
Qy	21	ArgGlyHisGluAspProGlyProLeuArgAlaLeuSerGlyGlnVal	36
Db	275	CGC-----CGCAGCCGCGAGTCGGGCGCGCGCGCGCGCGCGCGCGCGG	328
Qy	37	-----CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheVal	54
Db	329	CGCGCGTCCAGATATGCGGCGACGAGTCCGGGTGGGCTTCGACGGGGAGCCCTTCGTG	388
Qy	55	AlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArgGlu	74
Db	389	CGGTGCAACAGATGGCGCTTCCCTCTCTGCGCGCTCTACGAGTACGAGCGCGCGG	448
Qy	75	GlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArg	94
Db	449	GGCTCGCAAGCGTGCCTCGAGTGCAGACCGCTACAGCGCTCAAGGGCTGCCCGGG	508
Qy	95	VallalaGlyAspAspGluGluAspIleAspLeuGluHisGluPheAsnIleAsp	114
Db	509	GTGGCGCGCGCAGGAGGAGGAGCGGCTGCACGACCTGGAGGGCGAGTTCGGCTGCAG	568
Qy	115	AspGluAsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAla	134
Db	569	GACGGCGCGCCACGAGGAC-----GACCCGCGAGTACGTCCCGGAGTCC	613
Qy	135	MetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGlyAspGly---AsnAsn	153
Db	614	ATGCTCAGGGCGCAGATGAGTACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	673
Qy	154	ThrProGlnIleProProIleIleThrGlySerArgSerValProValSerGlyGlu---	172

Db	674	GTCCCAACAGTGCCTCCTCACCAACGCGCAGATGGTTGATGATCATCCCGCGGAGCAG	733
Qy	173	-----PhePro-----IleThrAsnGlyTyrGlyHisGlyGluValSerSerSer	187
Db	734	CACGCGCTCGTGGCGTCTACATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	778
Qy	188	LeuHisLysArgIleHisProTyrProValSerGluProGly-----SerAla	203
Db	779	-----AAGAGGATCCACCGCTCCTTTCGAGATCCCAACCTTCCAGTGCACCGAGA	832
Qy	204	LysTrpAspGluLysLysGlu-----ValSerTrpLysGlu	215
Db	833	TCCATGGACCGTCCAGGATCTGCCCGCTACGAGATATGCGAGTGGCGTGGCGGAGG	892
Qy	216	ArgMetAspAspTrpLysSerLysGlnGlyIleLeu-----GlyGlyGly	230
Db	893	AGATGGAGGCTGGNAGCAGACGAGCGCTGCAGCATGTGCAGGCGGAGGTGGC	952
Qy	231	AlaAspProGluAspMetAspAlaAspValAlaLeuAsnAspGluAlaArgGlnProLeu	250
Db	953	GGTGATGGGATGGCGAGTGCAGATCTGCCTTAATGATGAAGCTAGGCGAGCAATG	1012
Qy	251	SerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal	270
Db	1013	TCCAGAAAGTCCCTATATCAAGCGGAATTAATCCCTACAGATGATTATCGTTATC	1072
Qy	271	ArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAla	290
Db	1073	CGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1132
Qy	291	IleGlyLeuTrpLeuValSerIleCysGluIleTrpPheAlaIleSerTrpIleLeu	310
Db	1133	TTTGCAATGTGGCTCATATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTG	1192
Qy	311	AspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeu	330
Db	1193	GATCAGTTCCTCCAAAGTGGCTTCCCAATCGAGAGAGAGACTTACCTGGAGCCGTTGT	1252
Qy	331	ArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThr	350
Db	1253	AGTTTGCAGAGGAGGTCAACCTCTCAGCTTCTCCAAATCGACTTCTTGTTCAGTAG	1312
Qy	351	ValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAla	370
Db	1313	GTTCATCCCAAGAGAACTCCCTGTGTACAGCGAACAACACTGTCTTCCATCTCTTCT	1372
Qy	371	ValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeu	390
Db	1373	GTGATTTATCCGGTTGAGAGGTCTCTGCTGCTATGTTTCTGATGATGGTGTGCAATG	1432
Qy	391	ThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLys	410
Db	1433	ACGTTTGCAGCATGTCTGAAACATCTGAAATTCGAAAGAAATGGTTCCTTTCAGCAA	1492
Qy	411	LysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeu	430
Db	1493	AAGTTTAAATACGAGCTCGTGTCTCTGAGTGGTACTTCCAAACAGAGATAGACTAC	1552
Qy	431	LysAspLysValGlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGlu	450
Db	1553	AAAGCAACAGGTGCTGCTTCAATTTGAGGAGGAGGCGGATGAAAGAGAGATACGAG	1612
Qy	451	GluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGly	470
Db	1613	GAATTCAGGTAGATCAATGCTTGTGCAAAAGCCCAAAAGGTTCCTGTGAGAGAGA	1672
Qy	471	TrpIleMetLysAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGlyMet	490
Db	1673	TGGCAATGCAAGATGGAAGCCCTGCGCTGGAACCAACAGTACGCGATCATCTCTGGA	1732
Qy	491	IleGlnValPheLeuGlyHisSerGlyHisAspThrGluGlyAsnGluLeuProArg	510

```
Db 1733 ATTCAGGATTTCCTGGCCAAAGTGGCGTCTGATGTGGAAGAAATGATGTTGCCTCGC 1792
Qy 511 LeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAla 530
Db 1793 CTGGTTTATGCTCTCGAGAGAAAAGAGCCAGGTTATAACCATCAAGAAGAGCGTGGTGC 1852
Qy 531 MetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeu 550
Db 1853 ATGAATGACATGGTCCGCTGCTCTGCTGTCTTATCAAAATGCTGCATACCTATTGAACATTG 1912
Qy 551 AspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMet 570
Db 1913 GACTGTGATCACTACATCAACAAATAGCAAGGCCATAAAGAGGCTATGTGTTTCATGATG 1972
Qy 571 AspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIle 590
Db 1973 GATCCTTTGGTGGGGAAGAAAGTGTGCTATGTATGACAGTTCCCTCAGAGGTTTGATGGTATT 2032
Qy 591 AspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGly 610
Db 2033 GACAAAATGATCATAGCTACGCTAACAGGAACGTTGCTCTTTTGTGACATCAATGAAGGT 2092
Qy 611 LeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAla 630
Db 2093 TTGGACGGTATTCAAGGACCCATTATGTGGGTACTGGATGTGTTTTCAGACGGCAGGCA 2152
Qy 631 LeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysAspCys 650
Db 2153 CTGTATGTTATGATGCTCCTAAA---ACGAAGAAGCCACCATCAAGAATCTGCAACTGC 2209
Qy 651 CysPro-----CysPheGlyArgLysLysArgLys----- 660
Db 2210 TGGCCCAAGTGGTGCCTCTCTTGTCTGTGACGAGGAACAGAAATAAAGNAGACTACA 2269
Qy 660 ----- 660
Db 2270 AAACCAAGACGGAGNAGAAAGAAATTTTTCAGAAAGCAGAAACCCATCTCCT 2329
Qy 661 ---HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLys 679
Db 2330 GCATATGCTTTGGGTGAATTTGATGAAGGTCTCCA-----GGTGTGATATCGAGAAG 2383
Qy 680 GluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheVal 699
Db 2384 CCGGAAATCGTAATCAACAGAACTAGAGAAGAAATTTGGCGAGTCTCTGTTTGTCTG 2443
Qy 700 ThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAlaLeuLeu 719
Db 2444 GCATCAACACTTCTTGAGAACGGAGGACCCTGAAGAGCGCAAGTCCAGAGCTTCTCTTCTG 2503
Qy 720 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTyrGlyLeuGlu 739
Db 2504 AAGGAAGCTATACATGTTATCAGCTGCGGCTACGAAGAACAAGCCGACTGGGGAAGAGAG 2563
Qy 740 LeuGlyTyrPheLeuThrGluAspIleLeuThrGlyPheLysMetHisCys 759
Db 2564 ATTGGCTGGATTTACGGATCGATCAAGAGATATCTTGATGGATTTAAGATGCACATGC 2623
Qy 760 ArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaPro 779
Db 2624 CATGCTGGCGGTCTATTACTGTCATCCGAAGCGGCTGCATTCAAAGGTTCTGCGCT 2683
Qy 780 IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIle 799
Db 2684 CTGAACCTTTCCGACCGCTTTCACAGGCTCTTCTGCTGGCGCTTGGGTCCGTGGAATT 2743
Qy 800 PhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTyrLeu 819
Db 2744 TTCCTTCAGCAACACTGCGCCACTTTGGTACGGATAC---GGCGGCGGCTAAATTCCTG 2800
Qy 820 GluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAla 839
Db 2801 GAAAGGTTTCTTATATCAACTCCATCGTTTATCGCTGACGCTCCATTCCTCTCTCTGCTG 2860
```

```
Qy 840 TyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSer 859
Db 2861 TACTGTACCTTGGCTCCCATCTCCCTGCTCAGGGGAAGTTTATCACACAGAGCTTACC 2920
Qy 860 ThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeu 879
Db 2921 AATGTCCGCCAGTATCTGGTTTCATGGCACTTTTTCATCTGCATCTCCGTGACCGCATCCTG 2980
Qy 880 GluMetArgTyrSerGlyValSerIleGluGluTyrTyrArgAsnGlnPheTyrVal 899
Db 2981 GAAATGAGGTGGAGTGGCGTCGACATCGTGTGGAGGAACGAGCAGTTCGTGGTCC 3040
Qy 900 IleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAla 919
Db 3041 ATCGAGGCGTTTCGGCGCATCTGTTCCGGTGTTCAGGGCCTGCTGAAGGTGTTCCGCC 3100
Qy 920 GlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPhe 939
Db 3101 GGCATCGACACGAGCTTCACCGTGACGTCGAAGGCC---GGGGAC---GACGAGGAGTTC 3154
Qy 940 AlaGluLeuTyrAlaPheLysTyrThrThrLeuLeuIleProThrThrLeuLeuIle 959
Db 3155 TCGAGAGCTGTACACGTTCAAGTGGACACCCCTGCTGATACCCCGACCCAGCTCCTCTG 3214
Qy 960 IleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSer 979
Db 3215 CTGAACTTCATCGGGTGGTGGCGGGATCTCGAACCGGATCAACACGGGTACGAGTCG 3274
Qy 980 TrrGlyProLeuPheGlyLysLeuPhePheAlaPheTrrValIleValHisLeuTyrPro 999
Db 3275 TGGGGCCCCCTGTTTCGGGAAGCTCTTCTTCGCCCTTCTGGGTGATCGTCCACCTGTATCCCG 3334
Qy 1000 PheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIleTrrSer 1019
Db 3335 TTCTCTCAAGGCTGTGGTGGGAGGAGAACAGGACGCCGACGATCGTCATCGTCTGGTCC 3394
Qy 1020 IleLeuLeuAlaSerIlePheSerLeuLeuTrrValArgIleAspProPheIleValArg 1039
Db 3395 ATCTGCTGGCTCGATCTTCTGCTCTGTTGGTTCGCGCTCGACCCGTTCTCTCGCCAAG 3454
Qy 1040 ThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3455 AGCAACGGCCGCTCTCTGGAGGAGTGTGGCTGGAGTGC 3493
```

RESULT 6

```
US-10-160-719A-29
; Sequence 29, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321) ... (3551)
US-10-160-719A-29
```

Alignment Scores:

Pred. Name.:	0	Length:	3746
Score:	3912.50	Matches:	750
Percent Similarity:	78.96%	Conservative:	102
Best Local Similarity:	69.51%	Mismatches:	150
Query Match:	69.27%	Indels:	77
DB:	4	Gaps:	17

US-10-627-132-30 (1-1052) x US-10-160-719A-29 (1-3746)

Qy	34	GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe	53
Db	366	GGACAGGTGTCCAGATCTGGCGCCACGCGTGGGCACCAACGGCGGGGACGCTTTC	425
Qy	54	ValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArg	73
Db	426	GCCGCCCTCGCAGCTCTGGGGTTCCCGGTGTGCCGCCCTGCTACGAGTACGAGGC	485
Qy	74	GluGlyThrGlnAsnCysProGlnCysValysThrArgTyrLysArgLeuLysGlySerPro	93
Db	486	GACGGCAGCAGGGCTGCCCCAGTGCAGAACCAAGTACAGCGCCACACAGGGGAGCCCG	545
Qy	94	ArgValAlaGlyAspAspGluGluAspIleAspAspLeuGluHisGluPheAsn---	112
Db	546	CGCATCCGTGGG-----GAGGAGGACGACACTGATCGCATGGCAGCTTCAATTAC	599
Qy	113	IleAspAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThr	132
Db	600	CTTGCACTCTGGCAATGAGGAC-----CAGAAGCAGAAGATTGCC	638
Qy	133	GluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly-----	149
Db	639	GACAGAATGCCAGCTGCGCATGAACCTTGGGGCAGCGGGGATTTGGTCGCCCAAG	698
Qy	150	-----AspGlyAsnAsnThrProGln---Ile	157
Db	699	TATCACAGTGGCGAGATCGGCGCTTACCAGTATGACAGTGGGAGATTCCTCGGGGATAC	758
Qy	158	ProProlIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn	177
Db	759	ATCCATCATGTCACATAACAGCCAG-----ATCTCAGGAGAAATCCCTGGTGCTTCC	809
Qy	178	GlyTyrGlyHis---GlyGluValSerSerSerLeuHisLysArgIle---HisProTyr	195
Db	810	CCTGACCATCATATGATGTGCCCACTGGGAACATTTGGCAAGCGTCTCCATTTCCCTAT	869
Qy	196	ProValSerGluProGlySerAlaLysIleAspGluLysLys-----GluValSerTrp	213
Db	870	GTGAACCATTCGCCAAATCCGTCAGGGAGTTCTCTGGTACGATTGGGAATGTTCCCTGG	929
Qy	214	LysGluArgMetAspAspTrpLysSerLysGln-----GlyIleLeu-----	227
Db	930	AAAGAGAGGGTTGATGGCTGGAAATGAAGCAGGACAGGGGACGATTTCCCATGACGAAT	989
Qy	228	-----GlyGlyGlyAlaAspProGluAspMetAspAla-----	238
Db	990	GGCACAAAGCATTTGCTCCCTCTGAGGGTTCGGGGTTGGTGATATTGATGCATCAACTCAT	1049
Qy	239	-----AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysVal	254
Db	1050	TACAACATGGAAGATGCTTATTGAACGACGAAACTCGACAGCCTCTATCTAGAGAAAGTT	1109
Qy	255	SerIleAlaSerSerLysValAsnProTyrArgMetValIleValValArgLeuValVal	274
Db	1110	CCACTTCTCTCTCAGGATAAATCCATACAGGATGGTCATTTGCTCGGATTGATTGTT	1169
Qy	275	LeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrp	294
Db	1170	CTAAGCATCTCTTGCACTACCGTATCACAAATCCTGTGGCAATGCATACCCATTATGG	1229
Qy	295	LeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPhePro	314

Db	1230	CTTCTATCTGTTATATGAGATCTGGTTGGCTCTTTTCGTGGATATATGGATCAGTTCCCT	1238
Qy	315	LysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArg	334
Db	1290	AAGTGGTTTCCATCAACCGGAGACGACTCTGATAGGCTGGCATTAAGTATGACCGG	1349
Qy	335	GluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeu	354
Db	1350	GAAGGTGAGGCATCTCAGTTGGCTGCTTGTGACATTTTCGTAGTACAGTCGACCAATG	1409
Qy	355	LysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrPro	374
Db	1410	AAGGAGCCTCTCTTGTTCACGTGCAATACCGTCTATCCATTTCTCTGTGGATTACCT	1469
Qy	375	ValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSer	394
Db	1470	GTGGATAAGGTCTCTTGCATATCTATCTGATGATGGAGCTGCGATGCTGACATTTGATGCA	1529
Qy	395	LeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIle	414
Db	1530	CTAGCTGAGACTTTCAGAGTTTGTCTAGAAATGGGTACCATTTCTTAAAGATACACATT	1589
Qy	415	GluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysVal	434
Db	1590	GAACCTAGAGCTCCTGAATGGTACTTCTCCAGAAAATGATTACTTGAAGGCAAAAGTG	1649
Qy	435	GlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyrGluGluPheLysVal	454
Db	1650	CACCTTCATTTGTTAAAGACCCCGGGCCATGAGAGAGATATGAGAAATTCAAAGTT	1709
Qy	455	ArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLys	474
Db	1710	AGGGTAATGGCTTGTGTCTAAGGCACAGAAAGTTTCTTGAGGAAGGATGGATCATGCAA	1769
Qy	475	AspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPhe	494
Db	1770	GATGGCACACATGGCCAGGAAACAATAACAGGGACCATCTCGGAATGATTGAGGTTTTC	1829
Qy	495	LeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrVal	514
Db	1830	CTTGGTCACAGTGGTGGCTTGATCTAGGGGCATGAGCTACCCGTTTGGTCTATGTT	1889
Qy	515	SerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeu	534
Db	1890	TCTCGTAAAGCGTCTCGATTTCCAGCATCAACAGAAAGCTGGTGCCATGAATGCTCTT	1949
Qy	535	IleArgValSerIleValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHis	554
Db	1950	GTTGGTGTCTCAGCTGTGCTTACCAATGGACAATACTGTGAATCTTGATTTGTCATCAC	2009
Qy	555	TyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnVal	574
Db	2010	TACATTAACAAAGTAAGGCTCTCAGGGAGCTATGTGCTTCTTATGGACCTTAACCTA	2069
Qy	575	GlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAsp	594
Db	2070	GGAAGGAGTGTCTGCTAGCTCAGTTTCCCAGAGATTCGATGGCATTCACAGGAATGAT	2129
Qy	595	ArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIle	614
Db	2130	CGATATGCCAACAGGAACACCGTGTGTTTTTCGATATTTAACTTGAGAGGCTCTGATGCATC	2189
Qy	615	GlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyr	634
Db	2190	CAAGGACCAAGTTATGTGCGAATGGCTGTGTTTTCAACCGAACAGCTCTATATGGTTAT	2249
Qy	635	AsnProProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPhe	654
Db	2250	GAGCCCCCAATTAAAGCAGGAAGGGTGGTTCTTGTCACATCTATGTGGC	2300
Qy	655	GlyArgLysLys-----ArgLysHis-----ArgLysHis	661
Db	2301	GGTAGGAAGAACCAAGCAAGCAAAATCAAGAAAGGGCTCGGACAGAAAGTGTGCGAAGCAT	2360

```
QY 662 AlAlAspGlyLeuPro-----GluclyThrAlaAspMet 673
Db 2361 GTGGACAGTTCTGTGGCCAGTATTCAACCTTGAAGATATAGAGGAGGAGTTGAAGCGCT 2420
QY 674 GlyValAspSerAspGlyLeuMetLeuMetSerHisMetAenPheGluLysArgPheGly 693
Db 2421 GGATTGACGACGAGAAATCACTTCTTATGCTCAATGAGCCTGGAGAGAGATTGGC 2480
QY 694 GluSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSer 713
Db 2481 CAGTCCGCGAGCTTGTGCTCCACTCTGATGAGTATGCTGTGTCTCTCACTCCGCA 2540
QY 714 SerProAlaAlaLeuLeuLysGluAlaHisValIleSerCysGlyTyrGluAspLys 733
Db 2541 ACTCCGAGTCTCTTCTGAAGAGCTATCCATGTTATAGCTGTGGCTATGAGGACAG 2600
QY 734 ThrAspTrpGlyLeuGluLeuGlyTyrPheTyrGlySerIleThrGluAspIleLeuThr 753
Db 2601 ACTGAATGGGNACTGAGATCGGGTGGATCTACCGTTCTGTGACAGAGACATCTCACC 2660
QY 754 GlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAla 773
Db 2661 GGATTCAAGATGACGCGCGAGGCTGCGGTGCTACTTGCATGCCCAAGCGCCAGCT 2720
QY 774 PheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAla 793
Db 2721 TTCAAGGGTCTGCCCCCATCAATCTTTTCGACCGCTCTGAACCCAGGTCTCCGGTGGCT 2780
QY 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsn 813
Db 2781 CTTGGGTCGCGGAGATCCTCTTTCGCGGACCTGCGGCTGCTGCTGCTGCTGCTGCTGCT 2837
QY 814 GlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThr 833
Db 2838 GSGCGGCTCAAGTTCTCTGGAGAGATTGCGGTACATCAACACCATCTACCCGCTCAGC 2897
QY 834 SerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
Db 2898 TCCATCCCGCTTCTCATCTATCTGATCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2957
QY 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
Db 2958 ATCATTCAGAGATCAGCAACTTTCGCGAGCATCTGGTTTCATCTCCCTCTTCATCTCGATC 3017
QY 874 PheAlaThrGlyLeuLeuMetArgTrpSerGlyValSerIleGluGluTyrTrpArg 893
Db 3018 TTGCGCACGGGATCTCTGGAGATGAGTGGAGCGGGTGGGCGATCGACAGTGTGGAGG 3077
QY 894 AsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValGlnGly 913
Db 3078 AACGAGCAGTTCTGGGTGATCGGGGCGATCTCCGCGCACCTCTTCGCGCTGTTCCAGGGC 3137
QY 914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
Db 3138 CTGCTCAAGGTCTGCGCGGATCGACACCACTTCACCGTCACCTCCAGGCTCG--- 3194
QY 934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIlePro 953
Db 3195 GACGAGGACGGGACTTCCGGGAGCTGTACATGTTCAAGTGGACGACGCTCTCTGATCCG 3254
QY 954 ProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAlaIle 973
Db 3255 CCCACCATCTCTGATCATCAACTGCTGGCGCTGCTGCGCGGATCTCTCTACGCCATC 3314
QY 974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
Db 3315 AACAGGGATACAGTCTGGGGGCGGCTCTTCGCAAGCTCTTCTTGGCTTCTGGGTC 3374
QY 994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
Db 3375 ATCGTCCACTGATCCCGTTCTCAAGGGCTCATGGGCGGCTCATGGGCGAGACCGCACCCCGACC 3434
```

```
QY 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
Db 3435 ATCGTCTGTCTGGCCATCTCTGCTGCTCCATCTCTCTCTGCTGGGTTCGCATC 3494
QY 1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3495 GACCCCTTCACCACCGCGTCACTGCGCCGGATACCCAGACGTTGGCATCAACTGC 3551

RESULT 7
US-10-160-719A-49
; Sequence 49, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321) ... (3449)
US-10-160-719A-49

Alignment Scores:
Pred. No.: 0 Length: 3746
Score: 3912.50 Matches: 750
Percent Similarity: 78.96% Conservative: 102
Best Local Similarity: 69.51% Mismatches: 150
Query Match: 69.27% Indels: 77
DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-49 (1-3746)
QY 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53
Db 366 GGACAGGTGTCCAGATCTCGCGCACGGCTGGGCACCCACGCGGAGGGGACGCTCTC 425
QY 54 ValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArg 73
Db 426 GCGCGCTCGCAGCTCTCGGGTTTCGCGCTGCGCGCTCTGCTACGAGTACGAGCGGCA 485
QY 74 GluGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerPro 93
Db 486 GACGCGACGCGAGGCTGCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 545
QY 94 ArgValAlaGlyAspAspAspGluGluAspIleAspAspLeuGluHisGluPheAsn--- 112
Db 546 GCGATCCGTGGG-----GAGGAAGGAGACGACACTGATGCCGATAGCGACTTCAATTAC 599
QY 113 IleAspAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThr 132
Db 600 CTTGCATCTGCAATGAGGAC-----CAGAACGACGAGGAGGAGGAGGAGGAGGAGG 638
QY 133 GluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly----- 149
Db 639 GACAGATGCGCAGCTGGCGCATCAACGTTGGGGGCGAGCGGGGATGTTGGTCGCCCAAG 698
QY 150 -----AspGlyAsnAsnThrProGln-----Ile 157
```

Db 699 TATGACAGTGGCGAGATCGGGCTTACCAAGATATGACAGTGGCGAGATTCCTCGGGGATAC 758
Qy 158 ProProIleIleThrGlySerArgSerValProValSerGlyLeuPheProIleThrAsn 177
Db 759 ATCCCATCAGTCACTAACAGCCAG-----ATCTCAGGAGAAATCCCTGGTCTCC 809
Qy 178 GlyTyrGlyHis---GlyGluValSerSerSerLeuHisArgIle---HisProTyr 195
Db 810 CTGACCATCATATGATGTCCCAACTGGGAACATTTGGCAAGCGTCTCCATTTCCCTAT 869
Qy 196 ProValSerGluProGlySerAlaLysTrpAspGluLysLys-----GluValSerTrp 213
Db 870 GTGAACCATTCGCCAAATCCGTCAAGGAGTTCTCTGGTAGCATTTGGGATGTTCCTGG 929
Qy 214 LysGluArgMetAspAspTrpLysSerLysGln-----GlyIleLeu----- 227
Db 930 AAAGAGAGGGTTGATGGCTGGGAAATGAAGCAGGACCAAGGGGACGATCCCATGACGAAT 989
Qy 228 -----GlyGlyAlaAspProGluAspMetAspAla----- 238
Db 990 GGCACAGCATTTCTCCCTCTCAGGGTCCGGGTGGTGGTATTTGATGATCACTCAATGAT 1049
Qy 239 -----AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysVal 254
Db 1050 TACAACATGGAAGATGCCTTATTGAACGACGAACTCGACACCCCTCTATCTAGGAAGTT 1109
Qy 255 SerIleAlaSerSerLysValAsnProTyrArgMetValIleValValArgLeuValVal 274
Db 1110 CCATCTCTCTCCAGGATAAATCCATACAGGATGGTCAATGTGCTGCGATTGATTGT 1169
Qy 275 LeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrp 294
Db 1170 CTAAAGCATCTCTTCGACTACCGTATCAAAATCCTGTGCGCAATGCATACCCATATGG 1229
Qy 295 LeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPhePro 314
Db 1230 CTTCATCTGTATTATGTGAGATCTGGTTTGTCTCTTTCTGTGATATTTGGATCAGTTCCCT 1289
Qy 315 LysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArg 334
Db 1290 AAGTGGTTTCCAATCAACGGGAGACGTACTCTTGATAGCTGGCATTAAGGTATGACCGG 1349
Qy 335 GluGlyGluProSerLeuSerAlaValAspLeuPheValSerThrValAspProLeu 354
Db 1350 GAAGGTGAGCCATCTCAGTTGGCTGCTGTGTGACATTTTGTGTCAGTACAGTCGACCAATG 1409
Qy 355 LysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrPro 374
Db 1410 AAGGAGCTCTCTTGTCTACTGCCAATACCGTGTATCCATTTCTTGTGCTGTGGATTACCCT 1469
Qy 375 ValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSer 394
Db 1470 GTGGATAAGTCTCTTGTCTATGTAATGATGAGCTGGCATGCGCATTTTGATGCA 1529
Qy 395 LeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIle 414
Db 1530 CTAGCTGAGACTTCAAGATTGCTAGAAAATGGGTACCATTTTGTAAAGATGACAACTT 1589
Qy 415 GluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysVal 434
Db 1590 GNACTTAGACTCTCTGATGTTCTTCCCGAAATTTGATTACTTGAAGGACAAAGTG 1649
Qy 435 GlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysVal 454
Db 1650 CACCTTTCATTTGTTAAAGACCGCGGCCCATGAAGAGAGATAATGAAGATTCAAAAGT 1709
Qy 455 ArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLys 474
Db 1710 AGGGTAATGGCTTGTGTTGCTAAGGACAGAAAAGTTCTCTGAGGAAGGATGATCATGCAA 1769
Qy 475 AspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPhe 494
Db 1770 GATGACACACCATGGCCAGGAAACAAATACCAGGGACCATCTCTGGAAATGATTCAGGTTTC 1829

Qy 495 LeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrVal 514
Db 1830 CTTGTGTACAGTGGTGGCTTGATCTAGGGCAATAGTACCCGCTTTGTCTATGTT 1889
Qy 515 SerArgGluLysArgProGlyPheGlnHisIshisLysAlaGlyAlaMetAsnAlaLeu 534
Db 1890 TCTCGTGAAGAGCGTCTGGATTCAGCATCAAGAAGCTGGTGCCATGAATGCTCTT 1949
Qy 535 IleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHis 554
Db 1950 GTTCGTGTCTCAGCTGTCTTACCATGGACATACATGTTGAATCTTGATTTGTATCAC 2009
Qy 555 TyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnVal 574
Db 2010 TACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGTGCTTCTTATGGACCCTAACCTA 2069
Qy 575 GlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAsp 594
Db 2070 GGAAGGAGTGTCTGTCTAGTCCAGTTTCCACAGATTCGATGGCATTTGACAGGAATGAT 2129
Qy 595 ArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIle 614
Db 2130 CGATATGCCAACAGGACACCGTGTCTTTCATATTAATCTTGAGAGGCTTGTATGGCATC 2189
Qy 615 GlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyr 634
Db 2190 CAAGACCACTTATGTCCGAACCTGGCTGTGTTTCAACCGCAACAGCTCTATATGTTAT 2249
Qy 635 AsnProProLysGlyProLysArgProLysMetValThrCysAspCysProCysPhe 654
Db 2250 GAGCCCCCAATTAAGCAGAGAAGGGTGTGTTTCTTGTCA-----TCACTATGTGCG 2300
Qy 655 GlyArgLysLys-----ArgLysHis 661
Db 2301 GGTAGGAAGAAGGACCAAGCAATCAAGAAGGGCTCGGACAAAGAAGTCCGCAAGCAT 2360
Qy 662 AlalysAspGlyLeuPro-----GluGlyThrAlaAspMet 673
Db 2361 GTGGACAGTCTGTGTCAGTATTCAACCTTGAAGATATAGAGGAGGAGTTGAAGCGCT 2420
Qy 674 GlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGly 693
Db 2421 GGATTTGACGACGAGAAATCACTTCTTATGTCTCAATAGCTGGAGAGAGATTTGCG 2480
Qy 694 GlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSer 713
Db 2481 CAGTCCGCGAGCTTTGTTCCTCCACTCTGATGAGATATGTTGTGTCTCCTCAGTCCGA 2540
Qy 714 SerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLys 733
Db 2541 ACTCCGAGTCTCTTGAAGAAGCTATCATGTTATAAGCTGTGGCTATGAGGACAAG 2600
Qy 734 ThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThr 753
Db 2601 ACTGAATGGGAACTCAGATCGGTGGATCTACCGTCTCTGTGACAGAAGACATTTCCAC 2660
Qy 754 GlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAla 773
Db 2661 GGATTCAGATGACGCGCGAGGCTGGCGGTGATCTACTGATGCCCGCCAGGCGCAGCT 2720
Qy 774 PheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAla 793
Db 2721 TTCAAGGGGTCTGCCCCCATCAATCTTTGCGACCGTCTGAAACAGGTGCTCCGGTGGCT 2780
Qy 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrIleAsn 813
Db 2781 CTTGGGTCCGCGGAGATCTCTTTCAGCGGCACCTGCCCTGTGTGTACGGGTAC---GGA 2837
Qy 814 GlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrIleTyrProPheThr 833
Db 2838 GGGCGGCTCAAGTTCCTGGAGAGATTGCGTACATCAACACCATCTACCCGCTCAGG 2897

```
Qy 834 SerLeuProLeuLeuAlaTyrcysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
Db 2898 TCATCCGGCTTCTCATCTACTGATCCTGCCGCCATCTGTCTGTCCACGGGAAAGTTTC 2957
Qy 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
Db 2958 ATCATTCCAGAGATCAGCAACTTCGCCGAGCATCTGGTTTCATCTCCCTCTTCATCTCGATC 3017
Qy 874 PheAlaThrGlyIleLeuLeuMetArgTyrSerGlyValSerIleGluGluTyrTrpArg 893
Db 3018 TTCGCCACGGGCATCTCTGGAGATGAGTGGAGCGGGTGGGCATCGACGAGTGGTGAGG 3077
Qy 894 AsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGly 913
Db 3078 AACGAGCAGTCTCGGTGTATCGGGGCACTCTCCGGCAGCCTCTTCGGCGGTTCAGGGC 3137
Qy 914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
Db 3138 CTGCTCAAGGTGCTGGCGGCATCGACACCAACTTCACCGTCACTCCAGGGCTCG--- 3194
Qy 934 AspGluAspAspGluPheAlaGluLeuTyralaPheLysTrpThrThrLeuLeuIlePro 953
Db 3195 GACGAGGACGGCAGCTCTCGCGGAGCTGTATATGTTCAAGTGGACGAGCTCTCTGATCCG 3254
Qy 954 ProThrThrLeuLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAlaIle 973
Db 3255 CCCACCACTCTGTATCATCACTCACTGGTGGCGTCTGTGGCGGCACTCTCTACGCCATC 3314
Qy 974 AsnAsnGlyTyrcysGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
Db 3315 AACAGCGGATCACCAGTCTGTGGGGCCGCTCTTCGGCAAGCTCTTCTTCGGCTTCGGGTC 3374
Qy 994 IleValHisLeuTyrcysProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
Db 3375 ATCGTCCACCTGTACCCGCTTCTCAAGGGCTCTATGGCGAGCAACCGCACCCCGACC 3434
Qy 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033
Db 3435 ATCGTGTGTGTGGGCACTCTGTGGCGTCCATCTCTCTTCTGTGTGGTTCGCATC 3494
Qy 1034 AspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3495 GACCCCTTACACACCGCGCTCACTGGCCCGGATACCCAGACGCTGGGCATCAACTGC 3551

RESULT 8
US-10-160-719A-5
; Sequence 5, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-08-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338)...(3568)
US-10-160-719A-5
```

```
Alignment Scores:
Pred. No.: 0 Length: 3773
Score: 3912.50 Matches: 750
Percent Similarity: 78.96% Conservative: 102
Best Local Similarity: 69.51% Mismatches: 150
Query Match: 69.27% Indels: 77
DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-5 (1-3773)

Qy 34 GlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPhe 53
Db 383 GGAAGAGTGTGCCAGATCTCGCGCGACGGCGTGGGCAACACGGCGGGGGAGCGTCTTC 442
Qy 54 ValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgArg 73
Db 443 GCCGCTCTGCACGCTCTCGGGTTTCGGTGTGGCGCCCTGTCTACGATACGAGGCGCAAG 502
Qy 74 GluGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerPro 93
Db 503 GACGGCAGCGCAGCGCTGCCCGCAGTGCACAGCAAGTACCAAGCGCCACAGGGGAGCGCCG 562
Qy 94 ArgValAlaGlyAspAspAspGluGluAspIleAspAspLeuGluHisGluPheAsn--- 112
Db 563 GCGATCCGTGGG-----GAGGAAGGAGACGACACTGATGCCGATAGCGATCTCAATTAC 616
Qy 113 IleAspAspGluAsnGlnGlnArgGlnLeuGluGlnMetGlnAsnSerGlnIleThr 132
Db 617 CTGTCATCTGGCAATGAGGAC-----CAGAAGCAGAGAGATTGCC 655
Qy 133 GluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly----- 149
Db 656 GACAGAATGCGCAGCTGGCGCATGAACGTTGGGGCAGCGGGGATGTTGTGCGCCCAAG 715
Qy 150 -----AspGlyAsnAsnThrProGln-----Ile 157
Db 716 TATGACAGTGGCAGATCGGGCTTACCAAGTATGACAGTGGCGAGATTCTCGGGGATAC 775
Qy 158 ProProIleIleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsn 177
Db 776 ATCCCATCAGTCACTAACGCCAG-----ATCTCAGAGAGAAATCCCTGGTGTCTCC 826
Qy 178 GlyTyrGlyHis---GlyGluValSerSerSerLeuHisLysArgIle---HisProTyr 195
Db 827 CCTGACCATCATATGATGTCCCAACTGGGAACATTGGCAAGCGTGTCTCCATTTCCCTAT 886
Qy 196 ProValSerGluProGlySerAlaLysTrpAspGluLysLys-----GluValSerTrp 213
Db 887 GTGAACCATTTGCCAAATCCGTCAAGGGAGTCTCTCGTAGCATTTGGGAATGTTGCTCG 946
Qy 214 LysGluArgMetAspAspTrpLysSerLysGln-----GlyIleLeu----- 227
Db 947 AAGAGAGGGTGTGATGCTGGAAATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1006
Qy 228 -----GlyGlyGlyAlaAspProGluAspMetAspAla----- 238
Db 1007 GGCACAAGCATTTGCTCCTCTGAGGGTCCGGGTGTTGGTGATATATGATGATCACTAAT 1066
Qy 239 -----AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysVal 254
Db 1067 TACAACATGGAAGATGCTCTATTGAACGACGAACTCGACAGCCTCTATCTAGGAAGTT 1126
Qy 255 SerIleAlaSerSerLysValAsnProTyrArgMetValIleValValArgLeuValVal 274
Db 1127 CCATTCTCTCCAGGATAAATCCATACAGGATGGTCAATTGTGTGCGGATGATTGTT 1186
Qy 275 LeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrp 294
Db 1187 CTAAGCATCTTCTTGGCACTACCGGTATCACAATCTGTGCGCAATGATGATATATGATG 1246
Qy 295 LeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPhePro 314
Db 1247 CTTCTATCTGTATATGTGAGATCTGGTTTCTCTTCGTCGATATTTGGATCAGTTCCT 1306
```


QY 315 LysTrpPheProIleAspArgGluThrThrLeuAspArgLeuSerLeuArgTyrGluArg 334
DB 1307 AAGTGGTTCCTCAATCAACCGGAGACGTACTCTGTATGATGAGCTGGCATTAAGGTATGACCGG 1366
QY 335 GluGlyGluProSerLeuSerAlaValAspLeuPheValSerThrValAspProLeu 354
DB 1367 GAAGGTGACCAATCTCAGTTGGCTGTGTGATCTTTCTGTCAGTACGACCAATG 1426
QY 355 LysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrPro 374
DB 1427 AAGGAGCTCTCTGTCACTGCCAATACCGTGCTATCATCTTGTGCTGTGATACCT 1486
QY 375 ValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSer 394
DB 1487 GTGGATTAAGGTCTCTGTGCTATGTATCTGATGATGGAGCTGCATGTGATTCATGCA 1546
QY 395 LeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIle 414
DB 1547 CTAGCTGAGACTTCAGAGTTTCTAGAAATGGTACCATTTGTTAAGAAAGTACAACTT 1606
QY 415 GluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysVal 434
DB 1607 GAACCTAGAGCTCTGATGGTACTTCTCCAGAAATTTGATTACTTGAAGGACAAAGTG 1666
QY 435 GlnProThrPheValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysVal 454
DB 1667 CACCCCTTCATTTTAAAGACCGCGCGGCATGAAGAGAGAATATCAAGAATTCAAAGTT 1726
QY 455 ArgIleAsnAlaValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLys 474
DB 1727 AGGGTAATGGCTTGTGCTAAGGCACAGAAAGTTCTTGAGAGAGGATGGATCATGCAA 1786
QY 475 AspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPhe 494
DB 1787 GATGCACACCATGCGCAGGAACAATACACAGGACCATCTCTGGAAATGATTCAGGTTC 1846
QY 495 LeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrVal 514
DB 1847 CTTGGTCAAGTGGTGGCTTGATCTGAGGGCAATGAGTACCCCGCTTGGTCTATGTT 1906
QY 515 SerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeu 534
DB 1907 TCTGTGAAACGGCTCGGATTCAGCATCAAGAAAGCTGGTGCCATGAATGCTCTT 1966
QY 535 IleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHis 554
DB 1967 GTTCGTGCTCAGCTGTGCTTACCAATGGACAATACATGTGAATCTTGATGTGATCAC 2026
QY 555 TyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnVal 574
DB 2027 TACATTAACAACAGTAAGCTCTCAGGGAAGCTATGTGCTTCTTATGGACCCCTAACCTA 2086
QY 575 GlyArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAsp 594
DB 2087 GGAAGAGGTCTCTGCTAGCTCAGTTTCCCCAGAGATTGATGGCAATTCACAGGAATGAT 2146
QY 595 ArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIle 614
DB 2147 CGATATGCCAACAGAAACCCGTGTTTTCGATATTAATCTGAGAGGCTTGTATGGCATC 2206
QY 615 GlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyr 634
DB 2207 CAAGGACCACTTATGTGGAACTGGCTGTGTTTCAACCGGAACAGCTCTATATGTTAT 2266
QY 635 AsnProProLysGlyProLysArgProLysMetValThrCysAspCysProCysPhe 654
DB 2267 GAGCCCCCAATTAAGCAGAAAGGGTGTGTTCTTTGTCA-----TCACTATGTGGC 2317
QY 655 GlyArgLysLys-----ArgLysHis 661
DB 2318 GGTAGGAGAGGCAAGCAATCAAGAAGGGCTCGGCAAGAAAGTCCGAGAGCAT 2377

QY 662 AlaLysAspGlyLeuPro-----GluGlyThrAlaAspMet 673
DB 2378 GTGGACAGTTCGTCCAGATATTCAACCTTGAAGATATAGAGGAGGAGTTCAAGCGCT 2437
QY 674 GlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGly 693
DB 2438 GGAATTGACGACGAAATCACTTCTTATGCTCAATGAGCTCGAGAGAGATTTGGC 2497
QY 694 GlnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSer 713
DB 2498 CAGTCGCGAGCGTTTGTTCCTCAATGATGAGTATGGTGGTGTCTCTCAGTCCGCA 2557
QY 714 SerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLys 733
DB 2558 ACTCGGAGTCTCTCTGAAAAGACTATCCATGTTATAAGCTGTGGCTATGAGGACAG 2617
QY 734 ThrAspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThr 753
DB 2618 ACTGATGGGAACCTGAGATCGGTGGATCTACGGTCTGTGACAGAGACATTTCTCACC 2677
QY 754 GlyPheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAla 773
DB 2678 GGATTTCAAGATGCACGCGGAGGCTGGCGGTGATCTACTGATGCCCAAGCGGCAGCT 2737
QY 774 PheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAla 793
DB 2738 TTCAAGGGTCTGCCCCCATCAATCTTTCCGACCGTCTGAAACAGAGTGTCCGGTGGCT 2797
QY 794 LeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsn 813
DB 2798 CTTGGTCCGTGGAGATCCTCTTACGCGGACATGCCCCCTGTGGTACGGCTAC---GGA 2854
QY 814 GlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThr 833
DB 2855 GGGCGGCTCAAGTCTCTGAGAGATTCCGTGATCATCAACACCATCTACCCGCTCAGC 2914
QY 834 SerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPhe 853
DB 2915 TCCATCCGCTTCTCATCTACTGATCTGCGCGGCATCTGTCTGTCTCACCAGAAAGTTC 2974
QY 854 IleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIle 873
DB 2975 ATCATTCAGAGATCAGCAACTTCCGAGCATCTGTGTTCACTCCCTCTTCACTCTGATC 3034
QY 874 PheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArg 893
DB 3035 TTCGCCACGGCATCTCTGAGATGAGGTGAGCGGGTGGGCATCGACAGTGGTGGAGG 3094
QY 894 AsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValGlnGly 913
DB 3095 AACGAGCAGTTCGTGGTGATCGGGGGCATCTCCGCGCACCTCTTCGCGCTGTTCAGGCGC 3154
QY 914 LeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGly 933
DB 3155 CTGCTCAAGGTGCTGGCGGCGCATCGACACCAACTTCCACCGTCACTCCAGGCGCTCG--- 3211
QY 934 AspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIlePro 953
DB 3212 GACGAGGACGGCAGTCTCGCGGAGCTGTACATGTTCAAGTGGAGCAGCGCTCTGATCCCG 3271
QY 954 ProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyLysAspAlaIle 973
DB 3272 CCCACCACTCTGATCATCACTGTGTGGCGCTCGTCCGCGGCATCTCTTACGCCATC 3331
QY 974 AsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpVal 993
DB 3332 AACAGCGATACCACTGTGGGGCGGCTCTTCCGCAAGCTCTTCTCGCTCTTGGGTTC 3391
QY 994 IleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThr 1013
DB 3392 ATCTGCCACCTGTACCGTCTCTCAGGGCTCATGGCAGGACGACGACCCCGACCC 3451
QY 1014 ValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIle 1033

Db 3452 ATCTGCTGCTGCGGCATCTCTGCGGCCATCTCTCTCTGCTGCGGTTCCGATC 3511
Qy 1034 AspProPheLeuValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3512 GACCCCTTCAACCCCGCTCACTGCGCCGGATACCCAGACGTGTGGCATCAACTGC 3568

RESULT 9

US-10-160-719A-57
; Sequence 57, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10/160,719A
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272)...(3499)
US-10-160-719A-57

Alignment Scores:
Pred. No.: 0 Length: 3704
Score: 3903.50 Matches: 751
Percent Similarity: 78.41% Conservative: 99
Best Local Similarity: 69.28% Mismatches: 151
Query Match: 69.11% Indels: 83
DB: 4 Gaps: 19

US-10-627-132-30 (1-1052) x US-10-160-719A-57 (1-3704)

Qy 32 LeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAsp 51
Db 308 GTGGCGGCGCAGGTGTGCAGATCTGCGCGACGCGGTGGGCACCGCGCGGACGGCGAC 367
Qy 52 LeuPheValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGluTyrGlu 71
Db 368 CTCCTTCAACGCTGCGAGCTGCGCGCTTCCCGGTGCGCGCCATGCTACGATACGAG 427
Qy 72 ArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGly 91
Db 428 GCGAGGACGCGCAGCCAGCGGTGCGCGAGTGCAGACTAAGTACACGCGCCACAAAGGG 487
Qy 92 SerProArgValAlaGlyAspAspGluGlu---AspIleAspAspLeuGluHisGlu 110
Db 488 AGCCCAACAGTACACGGTGAGGAAATGAGGATGTGGATGTGTCAGAT----- 535
Qy 111 PheAsnIleAspAspGluAsnGlnGlnArgGlnLeuGluGlyAsnMet---GlnAsnSer 129
Db 536 -----GTGAGTGATACAACTAC-----CAAGCATCTGGCAACAGGATCAGAAGCAA 583
Qy 130 GlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGlyArgGlyProAspAspGly 149
Db 584 AAGATTGCTGAGAGATGCTCACTTGGCGGACAAAC---TCACGTGGCGAGTATATGGC 640
Qy 150 -----AspGlyAsnAsnThrPro 155
Db 641 CTGGCTAAGTATGACAGCGCGTGAATTTGGGCATGGGAAAGTATGACAGTGGTGAGATCCCT 700

Qy 156 Gln---IleProPheIleThrGlySerArgSerValProValSerGlyGluPhePro 174
Db 701 CGTGATATATCCCGTCACCTAACTCATAGCCAG-----ATCTCAGGAGGATTCCT 751
Qy 175 IleThrAsnGlyTyrGlyHisGlyGluValSerSer-----SerLeuHisLysArg 191
Db 752 -----GGAGCTTCCCTGATCATATGATGCTCTCTGTTGGGAACATTTGGCAGCGT 802
Qy 192 IleHisProTyrProVal-----SerGluProGlySerAlaLysTyrAspGlu 207
Db 803 GGACATCAATTTCTTATGTAATCATTTCTCCAAACCCATCGAGGGAGTTCTCCGGTAGC 862
Qy 208 LysLysGluValSerTyrLysGluArgMetAspTyrLysSerLys---GlnGlyIle 226
Db 863 CTGGCCATGTTGCATGGAAAGAGAGGGTGGATGGAAATGAAGGATAAAGGTGCA 922
Qy 227 Leu-----GlyGlyGlyAlaAspProGluAspMet 236
Db 923 ATTCCTATGACCAATGGAAACAAGCATTTGCTCCATCAGAGGGCGTGGAGTTGCTGATATT 982
Qy 237 AspAla-----AspValAlaLeuAsnAspGluAlaArgGlnPro 249
Db 983 GATGCTTCTACTGATTATAACATGGAGATGCGCTTACTGAATGATGAACCTCGCAACCT 1042
Qy 250 LeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetValIleVal 269
Db 1043 CTATCTAGAAAGTGCATTCCTTCATCCAGATAAATCCGTACAGAAATGGTCATTGTG 1102
Qy 270 ValArgLeuValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAsp 289
Db 1103 CTAGTTTGGTGTCTTATGCATATTTCTGCGTTACCGTATCACACATCTCTGTGAACAAT 1162
Qy 290 AlaIleGlyLeuTyrPheLeuValSerIleCysGluIleTyrPheAlaIleSerTyrIle 309
Db 1163 GCATATCCACTGTGGCTTTTATCCGTATATGTGAGATCTGGTTTGTCTCTGGATT 1222
Qy 310 LeuAspGlnPheProLysTyrPheProIleAspArgGluThrTyrLeuAspArgLeuSer 329
Db 1223 TTGGATCAGTTCCCAAGTGGTCCCAANTCAACCGTGAAACATACCTTGTAGACTGGCT 1282
Qy 330 LeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSer 349
Db 1283 TTAAGGTATACCGAGAAGGTGAACCATCTCAATTAGCTCTCTGTGTGATATTTTTGTGAGT 1342
Qy 350 ThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeu 369
Db 1343 ACTGTGGATCAATGAAGGAGCCTCTCTGTCTACTGCAATACTGTCTTTCATTCCTT 1402
Qy 370 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 389
Db 1403 GCTGTCCATTATCCGTTGACAAGGTATCTTGTCTATGTTTCGGATGTGGAGCTGTATG 1462
Qy 390 LeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTyrValProPheCys 409
Db 1463 CTGACTTTGATGCTCTCTGAAACTTCAGAGTTTGTAGAAAATGGTTCCGTTCTGT 1522
Qy 410 LysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyr 429
Db 1523 AAGAAGTACAACTAGAGCCTTAGGGCCCGGAATGGTACTTTGTCTCAGAAATTTGATTAC 1582
Qy 430 LeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArgGluTyr 449
Db 1583 TTGAAGACAAAGTTCAACCTCATTTGTGAAGAACCGCGCGCCATCAAGAGAGAAATAT 1642
Qy 450 GluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValProAlaGlu 469
Db 1643 GAAGAAATCAAAAGTTCGTATCAATGGTCTTTAGCAGGACCAAAAAGTTCCCGAGAG 1702
Qy 470 GlyTyrIleMetLysAspGlyThrProTyrProGlyValAsnAsnThrArgAspHisProGly 489
Db 1703 GGATGGATCATGCAAGATGGTATACCTTGGCCCTGGGAACAATACTTAGGAGCAATCTCTGGA 1762

QY 490 MetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGluLeuPro 509
Db 1763 ATGATTCAGGTTTCTCGGTCACAGTGGAGGCTTGACGTTGAAGGCAATGAATTCCT 1822
QY 510 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisIleLysLysAlaGly 529
Db 1823 CGTTTGGTTTATGTCTCTGTGAAACCGTCTCGATTCACACATCAACAGAGGCTGT 1882
QY 530 AlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsn 549
Db 1883 GCCATGAATGCATCTGTCTGTATCAGCTGCTTACTAATGGGCAATACATGTTGAAT 1942
QY 550 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCysPheLeu 569
Db 1943 CTGTGATGTGACCACTACATCAATATAGCAAGGCTCTTCGAGAAGCTATGTGCTTCCT 2002
QY 570 MetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnArgPheAspGly 589
Db 2003 ATGGACCCAAACCTAGGAAGGATGTCTGTATATGTCCTTCCATTTCTCAGAGGTTTGAATGT 2062
QY 590 IleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnMetLys 609
Db 2063 ATTGATAGGAATGACCGATATGCAACACAGGAACACTGTGTTTTCGATATTAACTTGAGA 2122
QY 610 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgGln 629
Db 2123 GGTCTTGACGGCATTCAGAGGGCAGTTTATGTGGAACTGGTGTGTGTTAAACAGAACG 2182
QY 630 AlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThrCysAsp 649
Db 2183 GCCTTATATGTTATGAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTCTCT 2236
QY 650 CysCysProCysPheGlyArgLysLys----- 658
Db 2237 ---TCGCTTTGTGGGGAGGAGAAAGACGCTCAAAATCTAAGAGAGCTCGSAAAGAG 2293
QY 659 -----ArgLysHisAlaLysAspGlyLeuPro-----Glu 668
Db 2294 AAGTCACATAGACACGACAGACTGTTCTGACCAAGTATTAAATCTCGAAGATATAGAGAA 2353
QY 669 GlyThrAlaAspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPhe 688
Db 2354 GGGATTGAAGGTTCTCAGTTTGATGATGAGAAATCGCTGATATGTTCTCAATGAGCTTG 2413
QY 689 GluLysArgPheGlyGlnSerAlaPheValThrSerThrLeuMetGluGluGly 708
Db 2414 GAGAAGAGATTGGCCAGTCCAGTGTGTTGTAGCTCTACTCTCATGGAATATGTTGT 2473
QY 709 ValProProSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCys 728
Db 2474 GTTCCACAATCTGCAACTCCAGAGTCTTCTGAAAGAGCTATTTCATGTCATCAGCTGT 2533
QY 729 GlyTyrGluAspLysThrAspTyrGlyLeuGluLeuGlyTyrIleTyrGlySerIleThr 748
Db 2534 GGCTATGAGGACAAACTGACTGGGAACTGAGATTGGTGGAATCTAGTTCTGTTACA 2593
QY 749 GluAspIleLeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMet 768
Db 2594 GAAGACATCTCACCGGATTCAGATGCATGCTCGAGGCTGGCGATCAATCTACTGCATG 2653
QY 769 ProllysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGln 788
Db 2654 CCTAAGCCACAGGCTTTCAAGGGATCTGCTCTATCAACCTTTCGGATCGTTTGAATCAA 2713
QY 789 ValLeuArgTyrAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeu 808
Db 2714 GTGCTTCGGTGGGCTCTTGGTTCCATGAAATCTTTTCAGCAGGCAATGCCATATGG 2773
QY 809 TyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThr 828
Db 2774 TATGGCTAT---GGAGGCGGCTTAATTCCTGGAGAGATTGCTTATATCAACACACA 2830
QY 829 IleTyrProPheThrSerLeuProLeuAlaTyrCysThrLeuProAlaValCysLeu 848

Db 2831 ATTTATCCACTCACATCAATCCGCTCCTCTCTACTCATATTGCCAGCAGTTTGTCTT 2890
QY 849 LeuThrGlyLysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAla 868
Db 2891 CTCACTGGAGAGTTTCATCATCCCAAGATTAGTAACCTAGAGAGTGTGGTTTATATCG 2950
QY 869 LeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValSerIle 888
Db 2951 CTCCTTATCTCAATCTTTGCCACTGGTATCTTGGAGATGAGTGAGTGGTGTGGCATT 3010
QY 889 GluGluTyrTrpArgAsnGluGlnPheTrpValIleGlyGlyValSerAlaHisLeuPhe 908
Db 3011 GATGAATGGTGGAGAACGACGAGCTTCGGGTCAATTTGGTGTATTTCTGCCGATTTATTT 3070
QY 909 AlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr 928
Db 3071 GCGGCTCTCCAGGGTCTCTGAGAGTGTGCTGTGTATCGACACAGGCTTCACTGTACCC 3130
QY 929 SerLysAlaThrGlyAspGluAspGluPheAlaGluLeuTyrAlaPheLysTrpThr 948
Db 3131 TCTAAGGCCACT---GACGAAGAAGGTGATTTTGGCGAGCTTACATGTTCAAGTGGACA 3187
QY 949 ThrLeuLeuIleProProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGly 968
Db 3188 ACGTCTTCATGCCCAACCACTATTATTGATCATCAACCTGGTGGGCTGGTCTGCTGCG 3247
QY 969 IleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhe 988
Db 3248 ATTTCTTACGCAATCAATAGCGGTACCGTACGATCAGGAGACCTCTTTTCGGAAGCTCTTC 3307
QY 989 PheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGln 1008
Db 3308 TTTGGTCTCTGGGTGATTGTCCACTGTACCTTCTCAAGGGCTCATGGGGAAGCAG 3367
QY 1009 AsnArgThrProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeu 1028
Db 3368 AACCGCACGCGCATGCTGCTGTTGCTGGGCTATCTCTTGGTTCGATCTTTTCCCTG 3427
QY 1029 LeuTrpValArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCys 1048
Db 3428 ATGTTGGTTCGTATCGATCCATTCACACCCGGGCTCACTGGCCCTTGATATCGGAATGT 3487
QY 1049 GlyIleAsnCys 1052
Db 3488 GGCATCAACTGC 3499

RESULT 10

US-09-221-013A-9
; Sequence 9, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (217)..(3411)

Db 2134 AGCTCTTTATCTAAGCTCTGTGGTGA-----TCAGAAAGAAAGATTCCAAAGCTAAG 2187
Qy 663 ----- 663
Db 2188 AAGAGTCGACAAAAGAAATCAGGACGACATCTGACTCAACTGTTCTCTGATTCAAC 2247
Qy 664 ---AspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspGlyGluMetLeu 682
Db 2248 CTCGATGACATAGAAGAGGAGTTGAAGGTGCTGGTTTGTGATGATCAAAAGCGCTCTTA 2307
Qy 683 MetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAlaPheValThrSerThr 702
Db 2308 ATGTCGCAATGAGCTGAGGAGCGATTGTGACAGCTGCTGTTTGTTCCTCTTACC 2367
Qy 703 LeuMetGluGluGlyValProSerSerSerProAlaAlaLeuLeuLysGluAla 722
Db 2368 CTAATGGAATGGTGGTCTCTCTTCAGCAACTCCAGAAACTTTCTCAAGAGGCT 2427
Qy 723 IleHisValIleSerCysGlyTyrgluAspLysThrAspTrpGlyLeuGluLeuGlyTrp 742
Db 2428 ATCCATGTCATTAGTTGGTTATGAGGATAAGTCAGATTGGGGAATGGAGATTGGATGG 2487
Qy 743 IleTyrglySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrp 762
Db 2488 ATCTATGGTCTGTGACAGAGATATTCTGACTGGTTCAAATGATGCCCGCTGGATGG 2547
Qy 763 ArgSerValTyrcysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeu 782
Db 2548 CGATCCATTACTGTCATGCTTAAGCTTCAGCTTTCAGGGTCTGCTCTCATCATCTT 2607
Qy 783 SerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePheSer 802
Db 2608 TCAGATCGTCTGAACCAAGTGTGAGTGGGCTTTAGGTTCAGTTTCAGATTCTCTTCAGT 2667
Qy 803 ArgHisSerProLeuLeuTyrglyTyrlsAsnGlyAsnLeuLysTrpLeuLysArgPhe 822
Db 2668 CGCATTTGCTTATGTTATGGTTTAC---AATGGAGCGCTAAATTTCTTGAGAGTTT 2724
Qy 823 AlaTyrlleAsnThrThrIleTyrglyProPheThrSerLeuProLeuLeuAlaTyrcysThr 842
Db 2725 GGTATGTGAACACCAACCATCTACCTATCACCTCCATCTCTCTCATGTATTGTACA 2784
Qy 843 LeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThrPheAla 862
Db 2785 TTGTAGCGGTGTGCTCTTCCACCAACAGATTATTATTCTTCAGATTAGTAACATTGCA 2844
Qy 863 SerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArg 882
Db 2845 AGTATATGTTCTGTCT 2904
Qy 883 TrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGly 902
Db 2905 TGGAGTGGGTAGGATAGACCAATGGTGGAGAAACGACGAGTTGGTTCATTGGTGA 2964
Qy 903 ValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAsp 922
Db 2965 GTATCCGCTCATTTATGCTGTGTTCAAGGTATCTCTCAAGTCTCTCAAGTCTCTCTCA 3024
Qy 923 ThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAlaGluLeu 942
Db 3025 ACAAACTTCACAGTTTACCTCAAAAGCTTCA---GATGAAGCGGAGCTTTGCTGAGTTC 3081
Qy 943 TyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeuLeuIleIleAsnVal 962
Db 3082 TACTTGTTCATATGACCAACACTTCTGATTCGCCCAACGACGCTCTCATTTGTAACCTTA 3141
Qy 963 IleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrglnSerTrpGlyPro 982
Db 3142 GTGGGAGTGTGTGACGAGTCTCTTATGCTATCAACAGTGGATACCAATCATGGGGACCA 3201
Qy 983 LeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrglyProPheLeuLys 1002
Db 3202 CTCCTTGGTAAGTTCTTTTCCCTTCTGGGTGATTGTTCACTTGTACCTTTCTCCTCAAG 3261

RESULT 11

US-10-160-719A-13

; Sequence 13, Application US/10160719A

; Patent No. 6803498

; GENERAL INFORMATION:

; APPLICANT: Dhugga, Kanwarpal S.

; APPLICANT: Helehtjaris, Timothy G.

; APPLICANT: Bowen, Benjamin A.

; APPLICANT: Wang, Xun

; TITLE OF INVENTION: Maize Cellulose Synthases and Uses

; TITLE OF INVENTION: Thereof

; FILE REFERENCE: 0864C

; CURRENT APPLICATION NUMBER: US/10/160,719A

; CURRENT FILING DATE: 2002-06-03

; PRIOR APPLICATION NUMBER: US 60/096,822

; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: US 09/371,383

; PRIOR FILING DATE: 1999-08-06

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 3725

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (179) ... (3400)

US-10-160-719A-13

Alignment Scores:

Pred. No.: 0 Length: 3725

Score: 3899.00 Matches: 738

Percent Similarity: 77.19% Conservative: 125

Best Local Similarity: 66.01% Mismatches: 145

Query Match: 69.03% Indels: 110

DB: 4 Gaps: 18

US-10-627-132-30 (1-1052) x US-10-160-719A-13 (1-3725)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 179 ATGGCGGCCAACAAAGGGATGGTGGCAGGCTCTCACAAACCCGACAGTGTGTCATGATC 238
Qy 21 ArgGlyHisGluAsp-----ProLysProLeuArgAlaLeuSerGlyGlnVal 36
Db 239 CGCCACGACGCGACGCGCTGTCCCGCTAAGCCACGAGAGTGCAGATGGGCGAGTTC 298
Qy 37 CysGluIleCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
Db 299 TGCCAGATTTGTGGCAGACACTGTTGGCGTTTCAGCCACTGCTGATGTCTTTGTCCTGC 358
Qy 57 AsnGluCysGlyPheProValCysArgProCysTyrglyTyrglyGluArgGluGlyThr 76
Db 359 AATGAGTGTGCTTCCCTGTCTGCCGCTTGTCTATGATGATGACGCGCAAGGAGGGAAC 418
Qy 77 GlnAsnCysProGlnCysLysThrArgTyrlsArgLeuLysGlySerProArgValAla 96
Db 419 CAATGCTCCCTCAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGATTCAT 478
Qy 97 GlyAspAspGluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGlu 116

Db 479 GGTGATGATGAGGAGAGATGTTGATGACCTGGACCAATGAATTCACACTATAAG----- 532
Qy 117 AsnGlnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeu 136
Db 533 -----CAGGCCAAT----- 541
Qy 137 HisGlyArgMetSerTyrGlyArgGlyPro-----AspAspGlyAsp 150
Db 542 -----GGNAGGGCCGACAGTGGCGAGCTTCAAGAGATGACCGCTGAT 583
Qy 151 GlyAsnAsnThrPro-----GlnIleProProIleIleThrGlySer 164
Db 584 CTGTCCTTCATCTGCTCGCATGACCCACACCATCGGATTCACAGCTTACAAAGTGGCAA 643
Qy 165 ArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGly----- 182
Db 644 CAG-----ATATCTGGAGAGATCCCTGATGCATCCCTGACCGCTCATCTTCATCCGC 694
Qy 183 GluValSerSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySer 202
Db 695 AGTCAACATCGAGCTATGTTGATCCAAAGCGTTCCAGTTCTCTGTGAGGATTTGTG----- 748
Qy 203 AlalysTrpAspGluLysLysGlu-----ValSerTrpLys 214
Db 749 -----GACCCCTCGAAGCACTTGAATTCCTATGGCTTAATAGTGTGACTGGGAAG 799
Qy 215 GluArgMetAspAspTrpLysSerLysGln----- 224
Db 800 GAAAGAGTTGAGAGCTGGAGGTTAAACAGGACAAAAATATGTTGCAAGTGACTAATAAA 859
Qy 225 -----GlyIleLeuGlyGlyAlaAspProGluLysAspMetAspAla 238
Db 860 TATCCAGAGGCTAGAGGAGACATGAGGGGACTGGCTCAATGGAGAGATATGCAA--- 916
Qy 239 AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSer 258
Db 917 -----ATGGTTGATGATGCAGCCCTACCTTTGAGCCGCAATTTGCTCA 967
Qy 259 SerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePhe 278
Db 968 AACCAAGCTCAACCTTTACCGATAGTAATATCTTCCTCTTATCATCTGCTGCTTCTTC 1027
Qy 279 LeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIle 298
Db 1028 TTCCAAATATGATATGATCATCCAGTGGTATGCTTATGAGATTTGCTGATGATTCGTT 1087
Qy 299 IleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePro 318
Db 1088 ATCTGTGAGGCTGCTGTTGGCTTGTCTGCTTCTAGATCATGTTCCCAAAATGGTATCCA 1147
Qy 319 IleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluPro 338
Db 1148 ATCAACCGTGACATATCTGCACAGGCTTGCATTTGAGGTATGATGAGAGGGAGAGCCA 1207
Qy 339 SerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProPro 358
Db 1208 TCACAGCTGGCTCCCATTTGATGCTTTGTCAGTACAGTGGATTCATTTGAAGAACCTCCA 1267
Qy 359 LeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysVal 378
Db 1268 CTGATCACAGCCAACTGTTTGTTCATCTTGTGTGGATTAACCTCTGTTGACAAAGTG 1327
Qy 379 SerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThr 398
Db 1328 TCATGCTATGTTCTGATGATGGCTCAGCTATGCTGACTTTTGTGACTCTCTCTGAAACT 1387
Qy 399 AlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAla 418
Db 1388 CCGCAATTTGCTAGAAAGTGGTTCCCTTTTGAAGAGACAAATATTGAACCAAGAGCT 1447
Qy 419 ProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPhe 438
Db -----

Db 1448 CCAGAAATTTTACTTTGCTCAAAAAATAGATTACCTGAAGGACAAAAATTCACACCTTCATTT 1507
Qy 439 ValGluGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAla 458
Db 1508 GTTAAGGAAGAAGACGAGCAATGAAGAGAGAGATATGAAGAATTCAAATTAAGATCAATGCC 1567
Qy 459 LeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPro 478
Db 1568 CTGTTGTTCCAAAGACACAGAAAGTGCCTGAAGAGGGGTGGACCATGGCTGATGGAACGTCT 1627
Qy 479 TrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSer 498
Db 1628 TGGCCTGGGAATAACCCCTAGGACCATCTCTGGCATGATTCCAGGTGTTCTTTGGGGCAGAT 1687
Qy 499 GlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLys 518
Db 1688 GGTGGGCTTGACACTGATGGAATGAATTAACACGCTCTTGTCTATGTTCTCTGCTGAAAG 1747
Qy 519 ArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSer 538
Db 1748 AGACCAAGCTTTCCAGCATCACAGAGGCTGGTGCATGAATGCACATGATTCGTGTATCT 1807
Qy 539 AlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsn 558
Db 1808 GCTGTGCTGCAAAATGGTGGCTATCTTCTCAATGGATTTGTGACCAATTTCTTCAATAGC 1867
Qy 559 SerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVal 578
Db 1868 AGCAAAAGCTCTTAGAAGCAATGTGCTTATGATGATCCAGCTCTAGGAGGAAACT 1927
Qy 579 CysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsn 598
Db 1928 TGTATGTACAAATTTCCAAAGATTTGATGGCATTGCTGCACGATCGATATGCTAAT 1987
Qy 599 ArgAsnThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProVal 618
Db 1988 AGGAACATAGTCTCTTTGATATCAACATGAAGGCTCTAGATGGCATTTCAGGGTCCAGTC 2047
Qy 619 TyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro----- 636
Db 2048 TATGTGGAAACAGATGCTGTTTCAATAGGAGGCTTTGATGATATGATCTCTGTTTGT 2107
Qy 637 ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArg 656
Db 2108 ACTGAAGCTGATCTGGAACCTTAACATTTGTTGTTAAAGAGTGTCTGT-----GGTAGA 2158
Qy 657 LysLysArgLysHisAlaLys----- 663
Db 2159 AGGAAGAGAAAGAACAAAGAGTTATATGATAGTCAAAAGCCGCTATTATGAAGAGAACAGAA 2218
Qy 664 -----AspGlyLeuProGluGlyThrAlaAspMetGlyVal 675
Db 2219 TCTTCAGCTCCCATCTTTAAATGATGAAGACATCAGAGGGGTATTGAA-----GGTTAT 2272
Qy 676 AspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyClnSer 695
Db 2273 GAGGATGAAGGTCAAGTCTTATCTCCAGAGGAAATTTGGAGAAACGCTGTTGGTCAGTCT 2332
Qy 696 AlaAlaPheValThrSerThrLeuMetGluGlyValProProSerSerSerPro 715
Db 2333 CCAATCTTTCATTCATCCACCTTTTATGACTCAAGGTGGCATATACCACCTTCAACAAACCCA 2392
Qy 716 AlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAsp 735
Db 2393 GCTTCTCTACTGAAGGAGCTATCCATGTTATACGCTGTGGGTAGGAGACAAACTGAA 2452
Qy 736 TrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPhe 755
Db 2453 TGGGGAAGAGATTTGGCTGGATCTATGTTTCAAGTTACAGAGGATATTTCTGACTGGGTTT 2512
Qy 756 LysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLys 775
Db 2513 AAAATGTCATGCAAGAGGCTGGCAATCAATCTACTGTCATGCCACCAACGACCTGTTTTCAG 2572

Db 800 GAAAGAGTTGAGAGCTGGAGGGTTAAACAGACGACAAAAATATGTTGCAAGTCACTAATAAA 859
Qy 225 -----GlyIleuGlyGlyAlaaspProGluAspMetAspAla 238
Db 860 TATCCAGAGGCTAGAGAGACATGGAGGGGAGCTGCTCAAAATGAGAGAGATATGCAA--- 916
Qy 239 AspValAlaLeuAenAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSer 258
Db 917 -----ATGGTTGATGATGACAGCTACCTTTGAGCGCATTTGGCCAAATTCCTCA 967
Qy 259 SerLysValAenProTyrArgMetValIleValValArgLeuValValLeuAlaPhePhe 278
Db 968 AACCAAGCTCAACCTTTACCGGATAGTAATCTCCGCTCTTATCATCTGCTGCTTCTTC 1027
Qy 279 LeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIle 298
Db 1028 TTCCAATATCGTATCAGTCATCCAGTCGTAATGCTTATGGATTGTGGCTAGTATCTGTT 1087
Qy 299 IleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePro 318
Db 1088 ATCTGTGAGGCTCGCTTTGCTGCTTCTAGATCAGTTCCTCCAAATGATATCCA 1147
Qy 319 IleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGluPro 338
Db 1148 ATCAACCGTGAGACATATCTCGACAGGCTTGAGGTATGATAGAGAGGAGAGCCA 1207
Qy 339 SerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProPro 358
Db 1208 TCACAGCTGGCTCCCATTTGATGCTTTGTCTAGTACAGTGGATCCATTGAAGAACTCCCA 1267
Qy 359 LeuValThrAlaAenThrValLeuSerIleLeuAlaValAspTyrProValAspLysVal 378
Db 1268 CTGATCACAGCCACACGTTTGTCTGCTGCTGATTAACCTGTTGACAAAGTG 1327
Qy 379 SerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThr 398
Db 1328 TCATGCTATGTTCTGATGATGGCTCAGCTATGCTGCTTTGAGTCTCTCTCGAAACT 1387
Qy 399 AlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAla 418
Db 1388 GCCGAATTTGCTAGAAATGGGTTCCCTTTTGTGAAGAGCAATATTGAACCAAGAGCT 1447
Qy 419 ProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPhe 438
Db 1448 CCAGAAATTTACTTTGCTCAAAAATAGATTACCTGAAGAGCAAAATTCACCTTCATT 1507
Qy 439 ValGlnGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAla 458
Db 1508 GTTAAGGAAAGACGAGCAATGAAGAGAGATATGAAGAATTCMAAATAAGAATCAATGCC 1567
Qy 459 LeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPro 478
Db 1568 CTTGTTGCCAAGCAGCAAAATGGCTGAAGAGGGGTGGACCATGGCTGATGGAACTGCT 1627
Qy 479 TrpProGlyAsnAenThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSer 498
Db 1628 TGGCCCTGGGAATTAACCTTAGGACCATCTGGCATGATTACAGTGTCTTGGGCGACAGT 1687
Qy 499 GlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLys 518
Db 1688 GGTGGGCTTGACACTGATGGAATGAATTAACCACTCTGTCTATGCTCTCTCGTGAAGAAG 1747
Qy 519 ArgProGlyPheGlnHisLysLysValAlaGlyAlaMetAsnAlaLeuIleArgValSer 538
Db 1748 AGACAGAGGCTTTCAGCATCACAGAAGGCTGTGTGAATGAATGCACTGATTCGTGTATCT 1807
Qy 539 AlaValLeuThrAsnAlaProPheMetLeuAenLeuAspCysAspHisTyrIleAsnAsn 558
Db 1808 GTGTGCTGACAAATGGTGCCTATCTCTCAATGTGGATTGTGACCATTAATCAATAGC 1867
Qy 559 SerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVal 578
Db 1868 AGCAAGAGCTCTTAGAAGAGCAATGTGCTTTCATGATGGATCCAGCTCTAGGAAGGAAACT 1927

Qy 579 CysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsn 598
Db 1928 TGTATTGTACATTTCCACAAAGATTTGATGGCATTTGACATTCACCATGATATGCTAAT 1987
Qy 599 ArgAsnThrValPhePheAspIleAenMetLysGlyLeuAspGlyIleGlnGlyProVal 618
Db 1988 AGGAACATAGTCTCTTTGATATCAACATGAAGGCTAGATGGCATTCAGGCTCCAGTC 2047
Qy 619 TyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAenPro----- 636
Db 2048 TATGTGGGAACAGGATGCTGTTTCAATAGCAGGCTTTGATGATATGATCTCTGTTTG 2107
Qy 637 ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArg 656
Db 2108 ACTGAAGCTGATCGNAACCTAACATTTGTTGTAAGAGCTGCTGT-----GGTAGA 2158
Qy 657 LysLysArgLysHisAlaLys----- 663
Db 2159 AGGAAGAAAGAAACAAGAGTTATATGGATAGTCAAAAGCCGTATTATGAAGAGACAGAA 2218
Qy 664 -----AspGlyLeuProGluGlyThrAlaAspMetGlyVal 675
Db 2219 TCTCAGCTCCCATCTTTAACATGGAGACATCGAGAGGGTATTGAA-----GGTTAT 2272
Qy 676 AspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSer 695
Db 2273 GAGGATGAAGAGTCAGTGTCTTATGTCAGAGGAAATTTGAGAAACGCTTTGGTCAGTCT 2332
Qy 696 AlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerSerPro 715
Db 2333 CCAATCTTCATTCATCCACTTTATGACTCAAGGTGGCATACCACTTCAACAAACCCA 2392
Qy 716 AlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAsp 735
Db 2393 GCTTCTCTACTGAAGAGCTATCCATGTTATCAGTGTGGTACGAGGACAAACCTGAA 2452
Qy 736 TrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPhe 755
Db 2453 TGGGAAAGAGATTGCTGATCTATGTTTACAGGATATTTCTGACTGGGTTT 2512
Qy 756 LysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLys 775
Db 2513 AAAATGATCAAGAGGCTGGCAATCACTACTGTCATGCCACCAACCTGTTTCAAG 2572
Qy 776 GlySerAlaProIleAenLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGly 795
Db 2573 GGTCTCGACCAATCAATCTTTCTGATCGTCTTATCAGGTGCTCCGTGGGCTCTTGGG 2632
Qy 796 SerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsn 815
Db 2633 TCAGTGGAAATTCGCTTAGCAGACATGTCCTATATGTTATGGCTAC-----AATGGGCGA 2689
Qy 816 LeuLysTrpLeuGluArgPheAlaTyrIleAenThrThrIleTyrProPheThrSerLeu 835
Db 2690 TTGAAGCTTTTGGAGAGGCTGGCTTACATTAACACCATTTGTTATCCAATCACATCTGTT 2749
Qy 836 ProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMet 855
Db 2750 CCGCTTATCGCTATGTTGCTTCTGCTATCTGCTCTTCTTACCAATAAATTTATCAT 2809
Qy 856 ProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAla 875
Db 2810 CCTGAGATTAGTAATATGCTGGAAATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTGCA 2869
Qy 876 ThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGlu 895
Db 2870 ACTGGTATTTGGAGCTCAGATGGAGTGGTGTGGCATTTGAAGATTGGTGAGAAATGAG 2929
Qy 896 GlnPheTrpValIleGlyGlyValSerAlaHisLeuPheAlaValAlaGlnGlyLeuLeu 915
Db 2930 CAGTTTTGGTTATTGGTGGCACTCTGCTCCCATCTCTTCGGGGTGTTCACAGGCTCTGCTG

359 LeuValThrAlaAanThrValLeuSerIleLeuAlaValAspTyrProValAspLysVal 378
Db CTGATCAGACCAACACTGTTTGTGTCATCTTCTGCTGGGATACCTCTGTGCAAAAGTG 1327
379 SerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThr 398
Db TCATGCTATGTTCTGATGATGGCTCAGCTATGCTGCTTTGAGTCTCTCTCGAAACT 1387
399 AlaGluPheAlaArgLysTyrValProPheCysLysLysPheGlyIleGluProArgAla 418
Db GCCGAAATTGCTAGAAAGTGGTTCCCTTTTGTGAAGACACAATAATTGAACCAAGAGCT 1447
419 ProGluPheTyrPheSerLeuLysValAspTyrIleLysAspLysValGlnProThrPhe 438
Db CCAGAAATTTACTTTGCTCAAAAATAGATTACCTGAAGGACAAAATTTCAACCTTCATTT 1507
439 ValGlnGluArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAla 458
Db GTTAAGGAAAGACGAGCAGATGAAGAGAGATATGAGNAATTCAAAATAAGAAATCAATGCC 1567
459 LeuValAlaLysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThrPro 478
Db CTTGTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGGTGGACCATGGCTGATGGAACTGCT 1627
479 TrpProGlyAanAanThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSer 498
Db TGGCTGGGAATAACCTTAGGGACCAATCTCGCATGATTTCAGGTGTTCTTGGGGCAGAGT 1687
499 GlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLys 518
Db GGTGGGCTTGACACTGATGGAATGAATTACCAGCTCTGTCTATGTCCTCGTGAAAG 1747
519 ArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSer 538
Db AGACAGGCTTTCAGCATCACAAAGAGCTGGTGCATGAATGCACATGATTCGTGTATCT 1807
539 AlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsn 558
Db GCTGTGTCGACAAATGGTGCTATCTCTCAATGGGATGTGACCATTAATCTCAATAGC 1867
559 SerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysVal 578
Db AGCAAGCTCTTAGAAGCAATGCTTTCATGATGCATCCAGCTCTAGGAAGGAAACT 1927
579 CysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsn 598
Db TGTATGTACAAATTTCCACAAAGATTTGATGGCATTGACTTGCACGATCGATATGCTAAT 1987
599 ArgAanThrValPheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProVal 618
Db AGGAACATAGTCTCTTTTGATATCAACATGAAGAGTCTAGATGGCAATTCAGGTCAGTC 2047
619 TyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro----- 636
Db TATGTGGGACAGGATGCTGTTCAATAGGCAGGCTTTGTATGGATATGATCCCTGTTTTG 2107
637 ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArg 656
Db ACTGAAGCTGATCTGGAACCTAACATTTGTTTGAAGAGCTGCTGT-----GSTAGA 2158
657 LysLysArgLysHisAlaLys----- 663
Db AGGAAGAGAAAGAACAAAGAGTTATATGGAATAGTCAACCGCTATTATGAAGAGAACAGAA 2218
664 -----AspGlyLeuProGluGlyThrAlaAspMetGlyVal 675
Db TCTTCAGCTCCCATCTTTTAACATGGACATCCGAGGAGGTATTGA-----GTTAT 2272
676 AspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSer 695
Db GAGGATGAAGGTCAGTGTATTATGTCCAGAGAGAAATTTGGAGAAACGCTTTTGGTCAGTCT 2332

696 AlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerSerPro 715
Db CCAATCTTATTCATTCACCTTATATGACTCAAGTGGCATACCACTTCAACAACCCCA 2392
716 AlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAsp 735
Db GCTTCTCTACTGAAGGAAGCTATCCATGTTTATCAGCTGTGGGTACGAGGACAAACTGAA 2452
736 TrpGlyLeuLeuLeuGlyTyrIleTyrGlySerIleThrGluAspIleLeuThrGlyPhe 755
Db TGGGAAAAGAGATATGGCTGATCTATGTTTCAGTTACAGAGGATATCTGACTGGGTTT 2512
756 LysMetHisCysArgGlyTyrArgSerValTyrCysMetProLysArgAlaAlaPheLys 775
Db AANAATGCATGCAAGAGCTGGCAATCAATCTACTGATGCCACCAACGACCTGTTTCAAG 2572
776 GlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGly 795
Db GGTTCGTGCACCAATCAATCTTCTGATCGCTTAAATCAGGTGCTCGGTGGGCTCTTGGG 2632
796 SerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsn 815
Db TCAGTGGAAATTCGTCTTAGCACACATGTCCTATATGATGGCTAC-----AATGGCGCA 2689
816 LeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeu 835
Db TTGAGCTTTTGGAGAGCTGGCTTACATTAACACCATGTTTATCCAAATCACATCTGTT 2749
836 ProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMet 855
Db CCGCTTATCGCTATTTGTGTGCTTCTGCTATCTGCTTCTTACCAATAAATTTATCATT 2809
856 ProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAla 875
Db CCTGAGATTTAGTAATATGCTGGAATGTTCTTCAATCTTCTTTTGGCTCAATTTTCGCA 2869
876 ThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGlu 895
Db ACTGTTATATGGAGCTCAGATGGAGTGGTGGCATTGAAGATTGGTGAGAAATGAG 2929
896 GlnPheTrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeu 915
Db CAGTTTTGGGTATTGTGGCACCCTCTGCCCATCTCTTCGGGTGTTTCCAGGGTCTGCTG 2989
916 LysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGlu 935
Db AAAGTGTGGCTGGATTGATACCACTTCACGTTACCTCAAGGCACTCT--GATGAG 3046
936 AspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThr 955
Db GATGGCAGCTTTGCTGAGCTATATGTTTCAAGTGGACCAAGTTGCTCATCCTCCGACC 3106
956 ThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsn 975
Db ACTGTTCTGTCTATTAACCTGGTCGAAATGGTGCAAGAAATTCGTATGCCATTAACAGC 3166
976 GlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleVal 995
Db GGCTACCACAACTCTGGGTCCGCTCTTTGGAAGAGCTGTTCTCTCGATCGGGTATCCTC 3226
996 HisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValVal 1015
Db CATCTTACCCTTCTCCTCAAGGTCCTCATGGCAGGAGCAACCCGACCAACCAATCTGC 3286
1016 ValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPro 1035
Db ATCGTTTGGTCCATCTCTTGGCTCTATCTTCTTCTTCTTCTTGGGTGAAGATCGATCCT 3346
1036 PheIleValArgThr---LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db TTCATCTCCCCGACACAGAAAGCTGCCCTTGGGGCAATGTGGTGTGAACCTGC 3400

RESULT 14


```
QY 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 517
Db 1693 AGTGGTGGGTCGACACTGATGGAAATAGATTACCACTCTTGTCTATGCTCTCGTGA 1752
QY 518 LysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAenAlaLeuLeuVal 537
Db 1753 AAGAGACAGGCTTTCAGCATCACAGAGGCTGGTGCATGAATGAATGGCTGATTGCTGTA 1812
QY 538 SerAlaValLeuThrAenAlaProPheMetLeuAenLeuAspCysAspHisTyrIleAsn 557
Db 1813 TCTGCTGTGTCGACAAATGGTGGCTCTCTTCAATGGATGGACCACTTCAAT 1872
QY 558 AsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
Db 1873 AGCAGCAAGCTCTTAGAGAAAGCAATGTCTTCATGATGATCGGCTCTAGGAGGAAA 1932
QY 578 ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAla 597
Db 1933 ACTTGTTATGTACAAATTCACACAGATTTGATGGCAATTCGATGGCAGATCGATATGCT 1992
QY 598 AsnArgAenThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
Db 1993 AATCGGAACATAGTTTCTTTGATATCAACATGAAGGCTCTGGATGGCATTTCAGGGTCCA 2052
QY 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro--- 636
Db 2053 GTTTACGTGGGAACAGATGCTGTTTCAATAGACAGGCTTTGATGATGATGATGATGATGAT 2112
QY 637 ---ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGly 655
Db 2113 TTGACTGAAGCTGATCGAGCCAAACATCTGTTATTAAGAGCTGCTGT-----GGT 2163
QY 656 ArgLysLysArgLysHisAlaLys----- 663
Db 2164 AGAAGGAGAGAGAGAGAGAGAGATATATGATGATGATGATGATGATGATGATGATGATGAT 2223
QY 664 -----AspGlyLeuProGluGlyThrAlaAspMetGly 674
Db 2224 GAATCTTCAGCTCCCATCTTCATATGGAACATCGAAGAGGGTATTGAA-----GGT 2277
QY 675 ValAspSerAspLysGluMetLeuMetSerHisMetAenPheGluLysArgPheGlyGln 694
Db 2278 TACGAGGATGAAGGTCAGTCTGTTATGTCACAGAGAGAAATTTGGAGAAACGCTTTGGTCAG 2337
QY 695 SerAlaAlaPheValThrSerThrLeuMetGluGluGlyGlyValProProSerSerSer 714
Db 2338 TCTCCTATTTTCATTGCAATCCACCTTTATGACACAAAGGTGGCATAACCACTTCAACAAAC 2397
QY 715 ProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThr 734
Db 2398 CCAGCTTCTCTACTAAGAGAGCTATCCATGTCATCATGTTGATGATGATGATGATGATGATGAT 2457
QY 735 AspTrpGlyLeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGly 754
Db 2458 GAATCGGGAAGAGATTTGGCTGGATCTATGTTTCAGTAACGGAGGATATTCTGACTGG 2517
QY 755 PheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPhe 774
Db 2518 TTTAAATGTCATGCAAGGGCTGGCAATCAATCTACTGTCATGCCACCAACCTTTGTTTC 2577
QY 775 LysGlySerAlaProIleAsnLeuSerAspArgLeuAenGlnValLeuArgTrpAlaLeu 794
Db 2578 AAGGGTTCTGCAACCAATCAATCTTCCGATGCTTATATCAGTGCTCGTTGGGCTCTT 2637
QY 795 GlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGly 814
Db 2638 GGGTCAGTGGAAATCTGCTTAGTAGACATTTGCTCTATCTGTTATGTTTAC---AATGGA 2694
QY 815 AsnLeuLysTrpLeuGluArgPheAlaTyrIleAenThrIleTyrProPheThrSer 834
Db 2695 CGATTGAAGCTTTTGGAGAGGCTGGCTTACATCAACACATTTATGATATCAATCAATCACC 2754
QY 835 LeuProLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIle 854
```

```
Db 2755 ATTCCGCTTATTGCTTATTGTTGCTTCCGCTATCTGCTCTTACCAATAAATTTATC 2814
QY 855 MetProSerIleSerThrPheAlaSerLeuPheIleAlaLeuPheMetSerIlePhe 874
Db 2815 ATTCCGAGATTAGCAATTAATGCTGGGATGTTCTTCACTTCTTTTCGCTCCATTTT 2874
QY 875 AlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsn 894
Db 2875 GCACCTGTTATATTGGAGCTTAGATGAGTGGTGGTGGCATTGAAGATTGGTGAGAAAT 2934
QY 895 GluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeu 914
Db 2935 GAGCAGTTTGGTATTATGGTGCACTCTGCCCATCTCTTCGACAGTGTTCAGGGGCTG 2994
QY 915 LeuLysValLeuAlaGlyIleAspThrAenPheThrValThrSerLysAlaThrGlyAsp 934
Db 2995 CTGAAGTGTGTGGCTGGGATTGATCAACCTCACAGTTACCTCAAGGCATCT---GAT 3051
QY 935 GluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProPro 954
Db 3052 GAGATGGCGACTTTGCTGAGCTATATGTTTCAAGTGGACCACTTGTCTCATTTCTCTCCG 3111
QY 955 ThrThrLeuLeuIleLeuAenValIleGlyValValAlaGlyIleSerAspAlaIleAsn 974
Db 3112 ACCACTGTTCTTGTCTTAACTTGGTGGAAATGGTGCAAGAAATTTCTTATGCCATTAAAC 3171
QY 975 AsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIle 994
Db 3172 AGTGGCTTACCAATCTCGGGTCCGCTCTTTGGAAAGCTGTCTCTCGATCTGGGTGATC 3231
QY 995 ValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAenArgThrProThrVal 1014
Db 3232 CTCCATCTCTACCCCTCTCTCAAGGCTCTCATGGGAAGGCAGAACCGCACCAACAATC 3291
QY 1015 ValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAsp 1034
Db 3292 GTCATTTCTGGTCCATCTCTTTCATCTATCTTCTCTTGTCTGTGGTGAAGATCGAT 3351
QY 1035 ProPheIleValArgThr---LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3352 CCTTTCATCTCCCGACACAGAAAGCTGCTTGGGGCAATGTGGCTCAACTGC 3408

RESULT 15
US-10-160-719A-53
; Sequence 53, Application US/10160719A
; Patent No. 6803498
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Theoreof
; FILE REFERENCE: 0864C
; CURRENT APPLICATION NUMBER: US/10160,719A
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 3753
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)...(3408)
US-10-160-719A-53

Alignment Scores:
```


Pred. No.: 0 Length: 3753
Score: 3895.50 Matches: 735
Percent Similarity: 76.94% Conservative: 126
Best Local Similarity: 65.68% Mismatches: 147
Query Match: 68.97% Indels: 111
DB: 4 Gaps: 17

US-10-627-132-30 (1-1052) x US-10-160-719A-53 (1-3753)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGlnLeuValLeuLeu 20
Db 184 ATGGCGGCCAACAGGGGATGGTGGCGGCTCGCAACCGCAACGAGTTCGTATGATC 243
QY 21 ArgGlyHisGluAsp-----ProLysProLeuArgAlaLeuSerGlyGlnVal 36
Db 244 CCCCACGACGGCGATGTGCGCGGCTCGGCTAAGACCCACAAAGAGTGGCAATGGACAGGTC 303
QY 37 CysGluLeuCysGlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCys 56
Db 304 TCCAGATTGGGGTACCTCTGGGTGTTTCAGCCACTGGGTGATCTTTTGGCTGC 363
QY 57 AsnGluCysGlyPheProValCysArgProCysTyrGluTyrGluArgGluGlyThr 76
Db 364 AATGAGTGTGCTTCCCTGTCTGCGCCCATGCTATGATGATGAGCGCAAGGAGGGAAC 423
QY 77 GlnAsnCysProGlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAla 96
Db 424 CAATGCTGCCCCAGTGCAGACTAGATACAAAGACAGCAAGAGGTAGCCCTCGAGTTCAT 483
QY 97 GlyAspAspGluGluAspLeuAspLeuGluHisGluPheAsnLeuAspGlu 116
Db 484 GGTGATGAGGATGAGGAAGATTGATGACCTAGACAAATGAATCAACTAC----- 534
QY 117 AsnGlnArgGlnLeuGluGlyAsnMetGlnAsnSerGlnLeuThrGluAlaMetLeu 136
Db 534 ----- 534
QY 137 HisGlyArgMetSerTyrGlyArgGlyPro-----AspAspGlyAsp 150
Db 535 -----AAGCAAGGAGTGGGAAGGCCAGAGTGGCAACTGCAAGGAGATGATCTGAT 588
QY 151 GlyAsnAsnThrPro-----GlnIleProIleIleThrGlySer 164
Db 589 CTGCTTTCATCTGTCGCCATGAGCCACATCATCGAATTCACGCTGACAGCGGTCAA 648
QY 165 ArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGly----- 182
Db 649 CAG-----ATATCTGGAGAGATCTCTGATGCTTCCCTGACCGTCATCTATCCGC 699
QY 183 GluValSerSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySer 202
Db 700 AGTCCAACATCAGCTATGTTGATCCAAGCGTCCCAAGTTCCTGTGAGGATTGTG----- 753
QY 203 AlaLysTrpAspGluLysLysGlu-----ValSerTrpLys 214
Db 754 -----GACCCCTCGAAGGACTTGAATTCCTATGGGCTTAATGATGTTGACTGGAAG 804
QY 215 GluArgMetAspAspTrpLysSerLysGln----- 224
Db 805 GAAAGATTGAGAGCTGGAGGTTAAACAGGCAAAATATGATGCAAGTGCATATAA 864
QY 225 -----GlyLeuGlyGlyGlyAlaAspProGluAspMetAsp 237
Db 865 TATCCAGAGGCTAGAGGAGGACATGGAGGCGACTGGCTCAAAATGGAGAGATATGCAA 924
QY 238 AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla 257
Db 925 -----ATGGTGTGATGATCAGCGCTACCTTTGAGCCGCTATCTGTCCTCAATTC 972
QY 258 SerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhe 277
Db 973 TCAACACAGCTCAACTTTACCGGGTAGTGATCATCTCCGCTTATCATCTCTGTGCTTC 1032

QY 278 PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSer 297
Db 1033 TTCTTCCAGTATCGTGTGTCAGTCATCCAGTGGTGTATGCTTATGGATTATGGCTAGTATCT 1092
QY 298 IleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhe 317
Db 1093 GTTATCTCGAGGCTCTGGTTTGGCTTCTTGGCTTCTAGATCAGTTCCTCAAAATGGTAT 1152
QY 318 ProIleAspArgGluThrTyrIleAspArgLeuSerLeuArgTyrGluArgGluGlyGlu 337
Db 1153 CCAATCAACCGTGAGACATATCTTGACAGGCTTGCTTGGTATGATAGAGAGGAGAG 1212
QY 338 ProSerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluPro 357
Db 1213 CCATCAGAGCTGGCTCCCATGATGCTCTCGTCAGTACAGTGGATCCATTTGAAGAACCT 1272
QY 358 ProLeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys 377
Db 1273 CCATGATCAGCCACACATGTTTGTCCATTTCTTCTGGGATACCTCTGTTGACAA 1332
QY 378 ValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGlu 397
Db 1333 GTGTCTATGCTATGTTCTGTGATGATGGTTCAGCTATGCTGCTTTTGGTCTCTCTCAGAA 1392
QY 398 ThrAlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArg 417
Db 1393 ACCGAGAAATTTGCTAGAAAGTGGTTCCTTTTGTAAAGAACCAATATTTGAACCAAGA 1452
QY 418 AlaProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThr 437
Db 1453 GCTCAGAAATTTTACTTTGCTCAAAATAATGATTTACCTTGAAGGACAAATTTCACTTCA 1512
QY 438 PheValGlnArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsn 457
Db 1513 TTTGTTAAGAAAGAGCGCAATGAAGAGGAGTATGAAGAAATTCAAAGTAAAGATCAAT 1572
QY 458 AlaLeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThr 477
Db 1573 GCCCTTGTGCAAGCAGACAGAAAGTGCCTGAAGAGGGGTGGACCATGGCTGATGAACT 1632
QY 478 ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 497
Db 1633 GCATGGCTGGGAATATCTTAGGACCATCTGTCATGATTCAGTGTCTTCTGGGGAC 1692
QY 498 SerGlyGlyHisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 517
Db 1693 AGTGGTGGCTCGACACTGATGGAATAGATTACCAAGCTTGTCTATGTCTCTCGTGA 1752
QY 518 LysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgVal 537
Db 1753 AAGAGACAGGCTTTTCCAGCATCACAAGAGGCTGGTGCATGAATGCGCTGATTCGTGTA 1812
QY 538 SerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsn 557
Db 1813 TCTGCTGGCTGACAAATGGTGCCTATCTTCAATGGGATGGACCATTTACTTCAAT 1872
QY 558 AsnSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLys 577
Db 1873 AGCAGAAAGCTCTTAGAAGACATGCTTTCATGATGATCCGGCTCTAGGAAGAA 1932
QY 578 ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAla 597
Db 1933 ACTTGTATGTACAAATTTCCACAGAGATTTGATGGCATTTGACTTGACGATGATGCT 1992
QY 598 AsnArgenThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyPro 617
Db 1993 AATCGGAACATAGTATTTCTTTGATATCAACATGAAGGCTCTGGATGGCATTCAGGGTCCA 2052
QY 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnPro--- 636
Db 2053 GTTTACGTGGGAAACAGGATGCTGTTTCAATAGACAGGCTTTGTATGGATGATCGATCTGTT 2112
QY 637 ---ProLysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGly 655

```
Db 2113 TTGACTGAAGCTGATCTGGAGCCAAACATGTGTATTAAAGAGCTCTGT-----GGT 2163
Qy 656 ArgLysLysArgLysHisAlaLys----- 663
Db 2164 AGAAGGAAGAAAAGAAACAGAGTTATATGATAGTATAGTCAAAGCCGTATTATTAAGAGAGAACA 2223
Qy 664 -----AspGlyLeuProGluGlyThrAlaAspMetGly 674
Db 2224 GAATCTTCAGCTCCCATCTTCAATATGAAGACATCGAAGAGGGTATTGAA-----GGT 2277
Qy 675 ValAspSerAspLysGluMetLeuMetSerHisMetAenPheHisLysArgPheGlyGln 694
Db 2278 TACGAGGATGAAGGTCAGTCTTATGTCCAGAGGAAATTTGGAGAAACGCTTTGGTCAG 2337
Qy 695 SerAlaAlaPheValThrSerThrLeuMetGluGlyValProProSerSerSer 714
Db 2338 TCTCTATTTTCATTTGCAATCCACCTTTATGACACAAGGTGGCATACCACCTTCAACAAAC 2397
Qy 715 ProAlaAlaLeuLysGluAlaLeuHisValIleSerCysGlyThrGluAspLysThr 734
Db 2398 CCAGCTTCTCTACTAAAGGAAGCTATCCATGTCATCATCAGTTGTGGATATGAGGACAAACT 2457
Qy 735 AspTrpGlyLeuGluLeuGlyTrpIleThrGlySerIleThrGluAspIleLeuThrGly 754
Db 2458 GAATGGGGAAGAGATTTGGCTGGATCTATGGTTCAGTAACGGAGGATATTCTGACTGGG 2517
Qy 755 PheLysMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPhe 774
Db 2518 TTTAAATGTCATGCAAGGGCTGGCAATCAATCAATCTACTGTCATGCCACCAACGACCTTTGTTTC 2577
Qy 775 LysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeu 794
Db 2578 AAGGTTCTGCAACCAATCAATCTTTCCGATCGTCTTAATCAGGTGCTCCGTTGGGCTTT 2637
Qy 795 GlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGly 814
Db 2638 GGCTCAGTGGAAATCTCTTAGTAGACATTTCTCTATCTGGTATGGTTAC--AATGGA 2694
Qy 815 AsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrIleTyrProPheThrSer 834
Db 2695 CGATTGAAGCTTTTGGAGAGGCTGGCTTACATCAACACTATTGTATATCAATCAATCACATCC 2754
Qy 835 LeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIle 854
Db 2755 ATTCCGCTTATTGCTATGTGTGCTCCGCTATCTCCCTCTTACCCTTACCATAAATTTATC 2814
Qy 855 MetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePhe 874
Db 2815 ATTCCTGAGATTAGCAATATGCTGGGATGTTCTTCATCTCTTTTCGCCCTCCATTTT 2874
Qy 875 AlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsn 894
Db 2875 GCCACTGGTATATTGGAGCTTAGAGTGGAGTGGTGTGGCATTTGAAGATTGGTGGAGAAAT 2934
Qy 895 GluGlnPheTrpValIleGlyGlyValSerAlaHisLeuPheAlaValGlnGlyLeu 914
Db 2935 GAGCAGTTTGGGTTATTGGTGGCACCTCTGCCCATCTCTTGCAGTGTTCAGGGTCTG 2994
Qy 915 LeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAsp 934
Db 2995 CTGAAAGTTTGGCTGGGATTGATACCAACTTCACAGTTACCTCAAGGCATCT--GAT 3051
Qy 935 GluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpTrpThrLeuLeuIleProPro 954
Db 3052 GAGGATGGCGACTTTGCTGAGCTATATGTGTTCAAAGTGGACCCAGTTTGTCTCATTCCTCCG 3111
Qy 955 ThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsn 974
Db 3112 ACCACTGTCTTGTATTAACTGGTGGAAATGGTGGCAGGAAATTTTATGCCATTAAAC 3171
Qy 975 AsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIle 994
```

```
Db 3172 AGTGGCTTACCAATCTCTGGGTCCGCTCTTTTGGAAAGCTGTCTTCTCGATCTGGGTGATC 3231
Qy 995 ValHisLeuTyrProPheLeuLysGlyLeuMetGlyValArgGlnAsnArgThrProThrVal 1014
Db 3232 CTCCTATCTTACCCCTTCTCTCAAGGGTCTCATGGGAAGGCAGAACCCGCACCAACAATC 3291
Qy 1015 ValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAsp 1034
Db 3292 GTCATTTGCTGGTCCATCTCTTTCATCTATCTTCTCTTGTGGTGAAGATCGAT 3351
Qy 1035 ProPheIleValArgThr---LysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3352 CCTTTCATCTCCCGCACAGAAAAGCTGTCTGCTTGGGGCAATGTGGCGTCAACTGC 3408
```

Search completed: March 12, 2005, 14:31:17

Job time : 450 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 12, 2005, 07:57:24 ; Search time 7251 Seconds
(without alignments)
7030.049 Million cell updates/sec

Title: US-10-627-132-30

Perfect score: 5648

Sequence: 1 MEASAGLVAGSHRNELVLI.....IDPFIVRTKGPDRQCGINC 1052

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO_sptol/US10627132/runat_04032005_113218_21146/app_query.fasta_1.1223
-DB=GenEmbl -QWMT=fastap -SUPFLX=rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTWMT=sptol -NORM=ext -HEAPSIZB=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10627132 @CGN 1.1 4417 @runat_04032005_113218_21146 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5648	100.0	3443	8 AY372246	AY372246 Zea mays
2	5347.5	94.7	3631	8 AK121170	AK121170 Oryza sat
3	4947	87.6	165909	8 AP005420	AP005420 Oryza sat
4	4947	87.6	187410	8 AP005579	AP005579 Oryza sat

5	4809	85.1	3277	8 AY095297	AY095297 Populus t
6	4784.5	84.7	3414	8 BT013780	BT013780 Lycopersi
7	4652	82.4	3081	8 BT004543	BT004543 Arabidops
8	4652	82.4	3355	8 AY139754	AY139754 Arabidops
9	4646	82.3	3081	8 AF088917	AF088917 Arabidops
10	4483	79.4	3911	8 AY639654	AY639654 Pinus rad
11	4477	79.3	3851	6 BD236020	BD236020 Materials
12	4247.5	75.2	106962	8 AC140546	AC140546 Medicago
13	4247.5	75.2	109777	2 AC150446	AC150446 Medicago
14	4224.5	74.8	7234	6 AX048245	AX048245 Sequence
15	4224.5	74.8	7234	8 AF091713	AF091713 Arabidops
16	4224.5	74.8	33563	8 AT11086	AT11086 Arabidops
17	4177.5	74.0	84322	8 AF004509	AF004509 Lotus cor
18	4131.5	73.1	4029	8 AK100877	AK100877 Oryza sat
19	4101	72.6	4127	8 AK121193	AK121193 Oryza sat
20	4078	72.2	4208	8 AK073561	AK073561 Oryza sat
21	4073.5	72.1	3968	8 AF200531	AF200531 Zea mays
22	4055	71.8	3773	8 AY196961	AY196961 Populus t
23	3997	70.8	3538	8 AF200530	AF200530 Zea mays
24	3995.5	70.7	3746	8 AK100914	AK100914 Oryza sat
25	3992	70.7	3643	8 AY643520	AY643520 Acacia ma
26	3982	70.5	3997	8 AY221087	AY221087 Solanum t
27	3974.5	70.4	3812	8 AF200532	AF200532 Zea mays
28	3966	70.2	3907	8 AY483152	AY483152 Hordeum v
29	3946	69.9	3723	8 AF150630	AF150630 Gossypium
30	3943	69.8	3991	8 AY573572	AY573572 Populus t
31	3934.5	69.7	3809	8 AY162180	AY162180 Populus t
32	3932	69.6	3229	8 BT002335	BT002335 Arabidops
33	3919.5	69.4	3954	8 AK072356	AK072356 Oryza sat
34	3914.5	69.3	3795	8 AF200533	AF200533 Zea mays
35	3914.5	69.3	3799	6 AX338680	AX338680 Sequence
36	3912.5	69.3	3745	8 AF200528	AF200528 Zea mays
37	3909	69.2	3682	8 AF027174	AF027174 Arabidops
38	3904	69.1	4282	8 AK069196	AK069196 Oryza sat
39	3903	69.1	3222	6 AX652952	AX652952 Sequence
40	3901.5	69.1	3264	6 AX653232	AX653232 Sequence
41	3901	69.1	3614	6 AR267559	AR267559 Sequence
42	3901	69.1	3614	6 AX030946	AX030946 Sequence
43	3901	69.1	3614	6 BD022678	BD022678 Manipulat
44	3899	69.0	3725	8 AF200526	AF200526 Zea mays
45	3897.5	69.0	3676	8 AF200529	AF200529 Zea mays

ALIGNMENTS

RESULT 1	AY372246	Zea mays	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
LOCUS	AY372246	Zea mays	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
DEFINITION	AY372246	Zea mays	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
ACCESSION	AY372246	Zea mays	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
VERSION	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
KEYWORDS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
SOURCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
ORGANISM	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
REFERENCE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
AUTHORS	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
TITLE	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA,	3443 bp	linear	PLN 01-SEP-2004
JOURNAL	AY372246.1	GI:38532103	cellulose synthase catalytic subunit 12 (CesA12) mRNA			

```
/mol_type="mRNA"  
/db_xref="taxon:4577"  
1..3443  
/gene="CesA12"  
92..3250  
/gene="CesA12"  
/note="CesA12; ZmCesA12"  
/codon_start=1  
/product="cellulose synthase catalytic subunit 12"  
/protein_id="AAR23312.1"  
/db_xref="GI:38532104"  
/translation="MEASGLVAGSHNRNBLVLRGHDPKPLRALRGQVCEICGDEV  
GLTVGDLFLVACNECGFPVPCVEYERREGTCPOCKTRKYKELKSGPRVAGDDDEE  
DIDLEHEFNDENQOLEGNMONSOITEAMLHGRMSYGRGPDGDGDNVTPQIPPI  
ITGSRSPVSGPEPTITNGYGHGEVSSSLKRIHPYPVSEPSAKWDEKESVKEBWD  
DWKSKQILGGADPEMDADVALNDEARQPLSRKVSIASSKNPYPMVIVRVLVLA  
FFURYLILHVPDAGVLSIICEIWFALISWLDQPPKFPFIDIRETYLDRLSRYER  
EGPSLSVDLFTSVTDPLKEPLVANTLSILAVDPVDKVSVCYSDGASMLTF  
ESLSETAEFARKWVFCCKFGIIPRAPEFYSUKVDYLDKQVQPTFYQERRAMKREYE  
EFKVINALVAKAMVPAEGWIMKDGTPWPNTRDHPGMIQVFLHSGHDTGNEHL  
PLVYVSREKRPQPHKKAGNNALIRVSANLTNPFLNLDCDHVNNSKAIRAM  
CFLMDPOGRKVCVQFPQRPDGLDVHDIRANRNTVFEDINMKGLDGIQGPVYVGKG  
VFRQALYGNPPKPGKPKWVTCDCPCFRKKRKHAKDGLPEGTADMGVDSKEML  
MSHNFERRFGQSAFTSTLMERGVPSPSPAALKKEAHHVISCYEDKTDWGLEL  
GWYISITEDILTGFQKHCGRWSYCMFKRAAPKGSAPINLSDRLNOLVRWALGSE  
IFSRHSPLLYGKGNLKWLERPAYINTTYPSTLPLAYCTLPVCLLTGKFIIMP  
SISTFASLFTALPMSIATGILEMRSGVSIEMWENQOPWITGGVSAHLFAVVOGL  
LKVLAGDTHFTVTSKATGBDEDFEALYAKWTTLLIPTTLLIINIVIGVAGISDA  
INNGYQSWGFLFGKLFPAFWIVHLYPFLGLMGRQNRNTPVVVWINSILLASIFSLW  
VRIDPFIVRTKGPVDRQGCNC"
```

ORIGIN

```
Alignment Scores:  
Pred. No.: 0 Length: 3443  
Score: 5648.00 Matches: 1052  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-627-132-30 (1-1052) x AY372246 (1-3443)  
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20  
Db 92 ATGGAGGCCAGCGCGGCTGTGGCGGCTCGCACACCGAAGAGAGTGTCTGATC 151  
Qy 21 ArgGlyHisGluAspProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40  
Db 152 CGGGGCCACGAGGACCCCAAGCGCTGCGGGCGCTGAGCGGGCAGGTGTGCGAGATATGC 211  
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60  
Db 212 GGCACGAGGTGCGGCTCACGGTGGACGGCGACCTCTTCGTGCGCTGCAACGAGTGGCGC 271  
Qy 61 PheProValCysArgProCysTyTGluTyArgArgGluGlyThrGlnAsnCysPro 80  
Db 272 TTCCCGGTGTCGGCCCTGTACGAGTACGAGCGCGGGAGGGACGCGAGAACTGCCCC 331  
Qy 81 GlnCysLysThrArgTyLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100  
Db 332 CAGTGCAGAGACGCGCTACAAGCGCTCAAGGGAGCCCGAGGGTTGCGGGGACGATGAC 391  
Qy 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120  
Db 392 GAGGAGGACATCGACGCTCGAGCACGAGTTCAACATCGACGACGAGAAATCAGCAGAGG 451  
Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140  
Db 452 CAGCTGGAGGGCAACATCGACGACGACGACATCAGCGAGGCGATGCTCGACGCGCAGGATG 511  
Qy 141 SerTyrglyArgglyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProfile 160  
Db 512 AGCTACGGGAGGGGGCCCCGACGACGCGGACGCGCAACACCCCGCAGATCCCGCCCATC 571
```

```
Qy 161 IleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrgly 180  
Db 572 ATCACCGGCTCCCGCTCCGTCGGTGAGCGGTGAGTTCCTCGATTACCAACGGGTATGCG 631  
Qy 181 HisGlyGluValSerSerSerLeuHisLysArgIleHisProTyroProValSerGluPro 200  
Db 632 CACGGCGAGGTCTCGTCTCTCCCTGACCAAGCGCATCATCCGTACCCCTGTGTCTGAGCCA 691  
Qy 201 GlySerAlaIleTyroAspGluLysLysGluValSerTyroLysGluArgMetAspAspTyr 220  
Db 692 GCGAGTGCCAAAGTGGGACGAGAAAGAAAGTGGAGGAGGAGAGTGGACGATCGG 751  
Qy 221 LysSerLysGlnGlyIleLeuGlyGlyGlyValAspProGluAspMetAspAlaAspVal 240  
Db 752 AAGTCCAAAGCAGCGGCATCTCGCGGGCGCGCGCATCCGAAACATGAGCGCGACGTG 811  
Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260  
Db 812 GCACCTGAACGAGCGAGCGAGCGCGCTCTCGAGGAAAGTGTCTGATCGCGTCCGAGCAAG 871  
Qy 261 ValAsnProTyroArgMetValIleValValArgLeuValValLeuAlaPheLeuArg 280  
Db 872 GTGAACCCGTACCGGATGGTATCGTGGTGGTCTCGTGTGTCTGCGCTTCTTCCTCCGG 931  
Qy 281 TyrArgIleLeuHisProValProAspAlaIleGlyLeuTyrLeuValSerIleIleCys 300  
Db 932 TACGATATCTGCACCCCGTCCCGGAGCCATCGGGCTGTGGCTGTCTCTCCATCATCTGC 991  
Qy 301 GluIleTyrPheAlaIleSerTyrIleLeuAspGlnPheProLysTyrPheProIleAsp 320  
Db 992 GAGATCTGGTTCGCCATCTCTCGATCTCTCGACAGTTCGCCAAGTGTCTCCCATCGAC 1051  
Qy 321 ArgGluThrTyroLeuAspArgLeuSerLeuArgTyroGluArgGlyGluProSerLeu 340  
Db 1052 CGCGAGACGTACCTCGACCGGCTCTCCCTCAGGTACGAGAGGAGGAGCGCTCGCTG 1111  
Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuVal 360  
Db 1112 CTGTGGCGGTGGACCTTCTGTGAGCACCGTGGACCCCGCTCAAGGAGCGCGCTGTGT 1171  
Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyroProValAspLysValSerCys 380  
Db 1172 ACCGCCAACACCGTGTCTCATCTCGCGGTAGACTACCCGCTGGGCAAGGTCTCTCTGC 1231  
Qy 381 TyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400  
Db 1232 TACGTCTCCGACGACGCGCGTGTGATGCTGACGCTTCTGAGTCTGTCTCGGAGACGCGCG 1291  
Qy 401 PheAlaArgLysTyrValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420  
Db 1292 TTCGGCGCAAGTGGGTGCCCTTCTGCAAGAGTTCGGCATCGAGCCCCCGCGCGAG 1351  
Qy 421 PheTyroPheSerLeuLysValAspTyroLysAspLysValGlnProThrPheValGln 440  
Db 1352 TTCTACTTCTCGTCAAGGTGCGACTACTCTCAAGGACAAAGGTGCGAGCCACCTCTCGTGC 1411  
Qy 441 GluArgArgAlaMetLysArgGluTyroGluGluPheLysValArgIleAsnAlaLeuVal 460  
Db 1412 GAGCGCGCGCCCATGAAGAGAGAGTATGAGGAGTTCAGAGTTCGGATCAAGCGCTGTGT 1471  
Qy 461 AlalysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyTyroThrProTyrPro 480  
Db 1472 GCCAAGGCCATGAAGTTCGGCAGAGGGGTGGATCATGAAGGACGCGACCGCTGGGCC 1531  
Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500  
Db 1532 GGGAAACAACACCCCGGACCAACCCCGCATGATCCAGGTGTCTCTGGGCGCACAGCGCGCGC 1591  
Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyroValSerArgGluLysArgPro 520  
Db 1592 CACGACACCGAGGGCAACGAGTCTGCCCGCTCTGTGTACGTCTCTCCGTGAGAGGCGCGC 1651
```

QY	521	GlyPheGlnHisHisValysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaVal	540	Db	2732	ATGCGGTGGAGCGGGTGGAGGATCGAGGAGTGGTGGAGAAACGAGCAGTTCTGGGTTCATC	2791
Db	1652	GGATTCCAGCACCAAGAGCGCGCCATGAACGCTCTGATTGCGGTCTCCGCCGTG	1711	QY	901	GlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGly	920
QY	541	LeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys	560	Db	2792	GGCGCGTGTCCGCGCATCTCTTCGCGCTCGTGCAGGCGCTGCTCAAGGTCTCTCCCGGG	2851
Db	1712	CTGACCAACGCGCCATTCACTCACTTGGACTGTGATCACTACATCAACACACAGCAAG	1771	QY	921	IleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspGluPheAla	940
QY	561	AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr	580	Db	2852	ATCGACACCAACTTCACCGTCCACTCCAGGCGCACCGCGCAGCAGGACGAGTTCGCC	2911
Db	1772	GCCATTCCGGAGGCGCATGCTCTCATGGACCTCAGGTCCGCGGAAGTCTGCTAC	1831	QY	941	GluLeuTyrAlaPheLysTyrThrThrLeuLeuIleProProThrThrLeuLeuIle	960
QY	581	ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn	600	Db	2912	GAGCTCTACGCTTCAAGTGGACACGCTCTCATCCCGCCACCACCGCTGCTCATCAT	2971
Db	1832	GTTTCAGTTCGCGCAGAGTTTCACCGCATCGACGTGCACGCCGATCGCTAACAGGAAC	1891	QY	961	AsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrp	980
QY	601	ThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrVal	620	Db	2972	AACGTCATCGCGCGTCTGGCGCGCATCTCCGACGCCATCAACACGGGTACCGTCTGG	3031
Db	1892	ACCGTCTTCTCGACATCAACATGAAGGGCTGGACGGCATCCAGGCGCGGTGTACGTC	1951	QY	981	GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe	1000
QY	621	GlyThrGlyCysValPheArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyPro	640	Db	3032	GGGCGCCCTCTTCGGCAAGCTCTTCTTCGCTTCTGGGTTCATCGTCCACCTCTACCGTTC	3091
Db	1952	GGGACAGGGTGGTTCGCGCGCAGCGGCTCTACGGCTACAAACCTCCCAAGGAGCC	2011	QY	1001	LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIleTrpSerIle	1020
QY	641	LysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLysLysArgLys	660	Db	3092	CTCAAGGGGCTCATGGGCGCCAGAACAGGACGCCACCGTGTGTTCATCTGGTTCATT	3151
Db	2012	AAGAGGCCAAGATGGTACCTGCGACTGCTGCGCGCTGCTTCGGCGCGCAAGAGCGAAA	2071	QY	1021	LeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIleValArgThr	1040
QY	661	HisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSerAspLysGlu	680	Db	3152	CTGCTGGCTCCATCTTCTCCCTGCTCTGGGTTCAGGATCGACCTTTCATCGTCAGACC	3211
Db	2072	CACGCCAAGGACGGCTCCGGAGGGCACCGCTGATATGGAGTAGATAGGCACAGAG	2131	QY	1041	LysGlyProAspValArgGlnCysGlyIleAsnCys	1052
QY	681	MetLeuMetSerHisMetAsnPheGlnLysArgPheGlyGlnSerAlaAlaPheValThr	700	Db	3212	AAGGCGCGGACGTCAGGACGTGGCATCAATTGC	3247
Db	2132	ATGCTCATGTCACCATCAACTTCAGAAAGCGGTTCGGGCGAGTCCGCGCGTTCGTACG	2191	RESULT 2			
QY	701	SerThrLeuMetGluGluGlyGlyValProProSerSerSerProAlaAlaLeuLys	720	AK121170	3631 bp	ORyza sativa (japonica cultivar-group)	linear
Db	2192	TCGACGCTGATGAGGAAGGCGCGTCCCTCTCTGTCGAGCGCGCGCTCCCTCAAG	2251	LOCUS			PLN 29-OCT-2003
QY	721	GluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTrpGlyLeuGluLeu	740	DEFINITION		insert sequence.	
Db	2252	GAGGCCATCCATGTCATGCTGGCTACGAGGACAGACGACTGGGGCTGGAGCTG	2311	ACCESSION		AK121170	
QY	741	GlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArg	760	VERSION		AK121170.1	GI:37990793
Db	2312	GGGTGGATCTACGGGTCCATCAGGAGGACATCCTGACGGGGTTCACAGTCACTGCGCG	2371	KEYWORDS		FLI_CDNA; CAP trapper.	
QY	761	GlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIle	780	SOURCE		ORyza sativa (japonica cultivar-group)	
Db	2372	GGGTGGCGCTCCGTGTACTGTGATGCCGAAGCGGGCGGTTCAGAGGGTTCGCGCGCATC	2431	ORGANISM		ORyza sativa (japonica cultivar-group)	
QY	781	AsnLeuSerAspArgLeuAsnGlnValLeuArgTTPAlaLeuGlySerValGluIlePhe	800	REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.	
Db	2432	AATCTATCGGACGCTCTCAACAGGTGCTCGGGTGGGCGCTGGGGTCCGTTCAGATCTTC	2491	AUTHORS		The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Koima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Onosuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Aiken, J., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kawaguchi, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.	
QY	801	PheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGlu	820	TITLE		japonica rice	
Db	2492	TTACGCGGCACAGCCCTCTGTATCGGCTACAGGACGCAACCTCAAGTGGCTGGAG	2551	JOURNAL		Science 301 (5631), 376-379 (2003)	
QY	821	ArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyr	840	MEDLINE		22752273	
Db	2552	CGCTTCGCGCTACATCAACACCATCTACCCCTTCACTCCGCTCCCGCTCGCTGCTAC	2611	PUBMED		12869764	
QY	841	CysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSerIleSerThr	860	REFERENCE		2	
Db	2612	TGCACCTCCCGCGCTGCTCTTACCGCGAAGTTTCATCTGCGCTCGATTAGCAG	2671	AUTHORS		Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,	
QY	861	PheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu	880				
Db	2672	TTGCGCAGCCTCTTCTTCATCGCCCTCTTCATGTCCTTCGCGAGCGGCATCTGGAG	2731				
QY	881	MetArgTrpSerGlyValSerIleGluTrpTrpArgAsnGluGlnPheTrpValIle	900				

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayaehida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, P., Hotta, I., Iida, Y., Ikeda, R., Imamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y.,
 Nakamura, M., Namiki, T., Narikawa, R., Niihara, J., Nishi, K.,
 Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M.,
 Oota, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C.,
 Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,
 Shibata, K., Shinagawa, A., Shiraki, T., Shiehiki, T., Sogabe, Y.,
 Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M.,
 Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K.,
 Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J.,
 Yokomizo, S. and Yoshimura, A.
 Collection, mapping, and annotation of 28K full-length cDNA clones
 from japonica rice

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
 3 (bases 1 to 3631)

Kikuchi, S.

Direct Submission

Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 32K full-length cDNA clones from japonica
 rice.

URL: <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T.,
 Yamamoto, M. and Nakahama, Y.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M.,
 Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M.,
 Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niihara, J., Oka, M.,
 Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M.,
 Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M.,
 Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
 Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

1. .3631
 /organism="Oryza sativa (japonica cultivar-group)"
 /moltype="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J023081B08"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3631

Score: 5347.50 Matches: 992

Percent Similarity: 96.78% Conservative: 30

Best Local Similarity: 93.94% Mismatches: 29
 Query Match: 94.68% Indels: 5
 DB: 8 Gaps: 4

US-10-627-132-30 (1-1052) x AK121170 (1-3631)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgHnGluLeuValLeu 20
 96 ATGAGGCGAGCGCGCGCTGGTGGCGGCTGCACACCGGAAACGAGTGTGTGTCATC 155
 QY 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluLeu 40
 156 CGGGGGCAGCGAGCCCAAGCCGCTGCGGGCCCTCAGCGGGCGAGTGTGCGAGATGC 215
 QY 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
 216 GGGCAGAGGTGGCGCGACCTGTCGAGCGGCGACTCTTGTGCTGCTGCAACGAGTGGCG 275
 QY 61 PheProValCysArgProCysTyrGluTyrGluArgGluGlyThrGlnAsnCysPro 80
 276 TTCGGGTGTGGCGCCCTGTCTACGAGTACGAGCGCGCGGAGGCCACCCAGAACTGCC 335
 QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
 336 CAGTGCAGACCCGCTACAGCGCCTCAAGGGGAGCCCGAGGGTGCAGGGGACGAGGAC 395
 QY 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGln 119
 396 GAGGAGACATTGACGACCTGGAGCAGCAGATTCAACATCGACCGAGAGCAGAGCAG 455
 QY 120 ---ArgGlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGly 138
 456 CTCGACGAGGATCAGGTGCGATGTCAGAACAGCACCATCACCGAGGGCGATGTCACGGC 515
 QY 139 ArgMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIlePro 158
 516 AAGATGAGCTACCGGAGGGGCCCGGACGACGCGGCGGCAACAGCACCCCG---CTCCG 572
 QY 159 ProfileThrGlySerArgSerValProValSerGlyGluPheProfileThrAsnGly 178
 573 CCGATCATCACCGCGCGTCTCGTCCGTCGCGGGGAGTTCCTCGATATCGAACAGC 632
 QY 179 TyrGlyHisGlyGluValSerSerSerLeuHisArgIleHisProTyrProValSer 198
 633 CATGGCCATGGCGAGTTCCTCTTCTCTGCACAAAGCGCATCCACCCCTTACCCGCTGCT 692
 QY 199 GluProGlySerAlaLysTyrAspGluLysLysGluValSerTyrLysGluArgMetAsp 218
 693 GAGCCAGGGAGTGCAAAGTGGGACGAGAAAGAGGTGAGTGGAAAGGAGAGGATGGAC 752
 QY 219 AspTyrLysSerLysGlnGlyValLeuGlyGlyValAla---AspProGluAspMetAsp 237
 753 GACTGGAAATCCAGACGAGGCGCATCGTCGCGGGCGGCGCCCGCATCCGACGATACGAC 812
 QY 238 AlaAspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAla 257
 813 GCCGACGTCCTCACTGAACGACGAGCGGCGAGCGCGCTGTTCGAGGAAGGTGTTCGATCG 872
 QY 258 SerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhe 277
 873 TCGAGCAAGGTGAACCCGTACCGATGTTGATCTCTCCGCTCTCGTGGTGTCTCGGCTTC 932
 QY 278 PheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTyrLeuValSer 297
 933 TTCTCTCGGTACCGCATCTCTCCACCGTGGCGAGCGCCATCCCGCTGTGGCTACCTCC 992
 QY 298 IleIleCysGluLeuTyrPheAlaIleSerTyrIleLeuAspGlnPheProLysTyrPhe 317
 993 ATCATCTGCGAGATCTGTTTCGCGCTGTCTCGTGGATCTCTCGACCATGTTCCCAAGTGTAC 1052
 QY 318 ProfileAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGlu 337
 1053 CCGATCGACAGGGGAGACCTACCTCGACCGGCTCTCTCCCTCGCTACGAGCGCGAGGGGAG 1112

AP005420	Oryza sativa (japonica cultivar-group)	165909 bp	DNA	linear	PLN 28-JUL-2004
LOCUS	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 9, PAC clone:P0418B08.				
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 9, PAC clone:P0418B08.				
ACCESSION	AP005420				
VERSION	AP005420.3				
KEYWORDS	GI:50725883				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group)				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.				
REFERENCE	1				
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.				
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC clone:P0418B08				
JOURNAL	Published Only in Database (2002)				
REFERENCE	2 (bases 1 to 165909)				
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-JUN-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan				
	(E-mail:tsasak@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)				
COMMENT	On Jul 27, 2004 this sequence version replaced gi:41152735. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/cdb/glimmerm/glmr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicesPredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0418B08 clone has an overlap with OJ1740.D06 (DDBJ: AP005579) clone at 5' end and with OJ1299.A11 (DDBJ: AP005568) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.				
FEATURES	Location/Qualifiers				
source	1..165909				
	/organism="Oryza sativa (japonica cultivar-group)"				
	/mol_type="Genomic DNA"				
	/cultivar="Nipponbare"				
	/db_xref="taxon:39947"				
	/chromosome="9"				
	/clone="P0418B08"				
	complement(5396..5638)				
gene	/gene="P0418B08.1"				
misc_feature	complement(5396..5638)				
	/gene="P0418B08.1"				
	/note="hypothetical ORF				
	predicted by GlimmerM				
	this category is not included in IRGSP standard"				
gene	25510..27251				
misc_feature	/gene="P0418B08.5"				
	25510..27251				

gene	predicted by GlimmerM				
mRNA	5856..10431				
	/gene="P0418B08.2"				
	join(5856..6221,6327..6562,6658..6754,6846..6972,7059..7671,7761..8024,8107..8319,8442..8649,8741..8937,9019..9372,9476..10431)				
	/gene="P0418B08.2"				
	/note="supported by full-length cDNA(s): AK121170"				
CDS	join(5951..6221,6327..6562,6658..6754,6846..6972,7059..7671,7761..8024,8107..8319,8442..8649,8741..8937,9019..9372,9476..10063)				
	/gene="P0418B08.2"				
	/note="contains EST(s): C72208(E1220), AU082309(E1220)				
	contains full-length cDNA(s): AK121170"				
	/codon_start=1				
	/product="putative cellulose synthase"				
	/protein_id="BAD33412.1"				
	/db_xref="GI:50725884"				
	/translation="MEASAGLVAGSHRNELVLIRGHEEPKPLRALSGQVCEICGDEV				
	GRIVDGLFVACNCEGFPVPCYECYERREGTQNCPCQKTRYKRLKGSRRVPGDEDEE				
	DIDDLLEHFNIDDEKQIQODQDGMQNSHITeamLHGKMSYGRGPDGDGNSGTPLPP				
	IITGARSVPVSGEPFISNSHGGEFSSSLHKRIHPYPVSEPGSAKDEKEVSKWMR				
	DWMKQGIIVAGADPDYDADVPLNDEARQPLSRKVSIASSKNVYRNVIILRLVY				
	LGFILRYLHPVPDAIPLMLT7SIICEIWFVSWILDQFPKWTPIIDRETYLRLSLRY				
	TPESLSETAFAFARVFCFKTSIEPRADEFYFSQKVDYLDKRVHFNFOERRAMKRE				
	YEEFKVIRNALVAKQKVAEGWIMKDGTPWGNTRDHPGMIQVFLHSGGHDTEGN				
	ELPRLVYSREKRGPGHKKAGAMNALLIRVSNVLTNPMLNDCDHVILNNSKAYRE				
	EUCFLMDPQVRKVCYVQPPQFDGLDHDVRYANRNTVFDINMGGLDGIQGVYGT				
	GCVFRQALYGNPKPKPKVMTCCPCFGRKKRGKGLGSLPAVADGMDSDK				
	EMLMSQNFKEFGQSAAVFTSLMBEGGVPPSSPAALKEAIIHVSCEYEDTDMG				
	LSELGTIYGSITDILGPKMCHGRWSVCMKRAAFKGSAPINLSDRLNQLRWALG				
	VEIIPESRPILLYGKNGLKWLEFSVINTIYPTSLPILAYCTLPVACLITGKF				
	IMPTTIFASLFFIALFISIFATGILEMRWSGVSIEWRNEQFWIGGSAHLFAVY				
	QGLLKVLADGNTFTVTSKATGDEDEFAELYAFKWTLLIPTTILLILNLIGVAGV				
	SDAINNGSEAWFLGKLFJFAFVIVHLYPFLKGLMGQRNRTTIVIVMSVLLASIFS				
	LWVRIDPTTIKARGPDVRQCGINC"				
	complement(10866..17118)				
gene	/gene="P0418B08.3"				
mRNA	complement(join(10866..11328,11488..11595,12606..12695,12806..12922,13597..13716,14119..15432,16817..17118))				
	/gene="P0418B08.3"				
	/note="supported by full-length cDNA(s): AK100679"				
CDS	complement(join(11263..11328,11488..11595,12606..12695,12806..12922,13597..13716,14119..15357))				
	/gene="P0418B08.3"				
	/note="contains full-length cDNA(s): AK100679"				
	/codon_start=1				
	/product="streptococcal hemagglutinin-like protein"				
	/protein_id="BAD33413.1"				
	/db_xref="GI:50725885"				
	/translation="MDVLKSDVKTKTGLNETLPPVLVPSSEKHNASPVNRRGDRVSRVK				
	NGLSAAATTAARCTSPGRTSANECPKRAQASADRRRSTPSRSTSTPSTPAS				
	RSYTPVNTVTEGHSRRITRTNDGLMPARNLSSSFQSSVVTTPGNKKDKVRVS				
	GLDQTKGQASVIAERKSRPLRNKIQECENAOQPSBEDQPRVIEQHRWPMQSGRVA				
	SNLTSIDMSDKAGSRVPSNTISRGSVPKTLASCTGKGNKSLDEVARRLAHAG				
	GRDQVDSRCHYQSQTERCKSVSRPRAVTLPPVLHRSRSPSKASVSTSSRSFQ				
	SPSRTPSPSRQSAGSIQSGVASPIISVMVDKKGKNNSQIENIHLHLSVNYRL				
	QKIFVNAVADETMSFQKVTAESIIYNVLRNTSLNRDVTNNRRINVCIIQELKLGIL				
	KEQIDYLEQWPALEKENSILSFHATEALKASTLRLPVTSGAKADVVALKNAVSAVDI				
	MQGLGSAVRCKLPKVEDRTYLVSELSVIARQEKAMLDCECELLAAKIQVQESSLRT				
	HLTLRPGIAHMI"				
gene	18866..19216				
misc_feature	/gene="P0418B08.4"				
	18866..19216				
	/gene="P0418B08.4"				
	/note="hypothetical ORF				
	predicted by GlimmerM				
	this category is not included in IRGSP standard"				
gene	25510..27251				
misc_feature	/gene="P0418B08.5"				
	25510..27251				

```
/gene="P0418B08.5"
/notes="probably inactive due to including stop codon(s) in
CDS
pseudogene, transposase"
/pseudo
complement (29005..36438)
/gene="P0418B08.6"
complement (join(29005..29109,29221..29352,29652..29818,
30448..30637,31719..31850,32464..32571,32702..32797,
32891..32998,33106..33278,34215..34590,35076..35219,
35314..35988,36106..36171,36271..36438))
/gene="P0418B08.6"
/notes="start and end point are not identified"
complement (join(29005..29109,29221..29352,29652..29818,
30448..30637,31719..31850,32464..32571,32702..32797,
32891..32998,33106..33278,34215..34590,35076..35219,
35314..35988,36106..36171,36271..36438))
/gene="P0418B08.6"
/notes="similar to Oryza sativa chromosome 7,
OSUNB006L20.19"
/codon_start=1
/product="hypothetical protein"
/db_xref="GI:50725886"
/translation="WASDRSSDDECSNSKDNKVDADSYANSISEEETSDDKYI
TKDLVSLDKLFGSKIRKARTDKPVRKQKRTKLEVDQDKSGSAESPSPRLN
ITYFNSIEGLSNQRRIENSFGSLNFORCAIPLSFVKWIASHTDVSQDIWVG
RSIPNNTNFILIGIPNGGLEIKNDNDAGKHFFHQFGSTKPLISFGTKLLSDKG
NKLSDDVRCFVVALSTFLCPNSDTHPSKYLEPLSDIKSSKWNWSKFVYEWLMT
YIAFORESKSEKTSKTLGCGCHLLAINVLDKDFGRNIPIGPPRISVWKGMIK
EYKIDCKTFQKGRPLDIDISTVCKWFRHSDDTYTGKNAQAARVLDVCSNKSAS
LFTQKIDTTREKSGEKSCKGSGKNVNAHNGNAPFLQNHPPNNDVCSQLPKP
QTSKTAHKTNNNTSTSGAPPTDKASSKDFVCSQSNNSVASRTRAKRLHSAS
PLSVISPSKYSDDKIRKARQKPLKATEPINTOPDREESNKDALFVEPICIIPA
KKEVQTKNLESSTFVIDIEGYDAEDITGHTDTKTFILVYNSSEEHNSODP
TODSDNIPNKSTNQPECLHFDKSKNSASLGHFLSQTSGTNIDIEHSSNQ
SSQSPDIGNSPRIAQNRPNQHAQEROYSMIRIISLNASAGNETHRNYRFPK
RIVHPILMDNVKVTWSSLSLSPGVVDVTVLNAKKIANDQNKENEYRNFYFH
RTSLIFFLMDNVWFVFTVDIKGHFVFLDSYDNNKNNKHKIQGLLIPGFIAMWEEF
SYVEKNFSKFDIQYPPITRQNG"
40242..41638
/gene="P0418B08.7"
join(<40242..40308,40655..40900,40996..41112,41239..41339,
41570..>41638)
/notes="start and end point are not identified"
join(40242..40308,40655..40900,40996..41112,41239..41339,
41570..41638)
/gene="P0418B08.7"
/notes="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/db_xref="GI:50725887"
/translation="MPATEELEATATWELDEAGEKAWGGALPHPPHPKPTAAVATRR
```

Alignment Scores:

Pred. No.:	0	Length:	165909
Score:	4947.00	Matches:	988
Percent Similarity:	74.18%	Conservative:	32
Best Local Similarity:	71.85%	Mismatches:	31
Query Match:	87.59%	Indels:	328
DB:	8	Gaps:	13

US-10-627-132-30 (1-1052) x AP005420 (1-165909)

Qy	1	MetGluAserAlaGlyLeuValAlaGlySerHisAenArgAenGluLeuValLeuile	20
Db	5951	ATGAGGCGAGCGCGGCTGGTGGCGGTCGACCAACCGGAACGAGCTGGTCTGATC	6010
Qy	21	ArgGlyHisGluAapProLysProLeuArgAlaLeuSerGlyGlnValCysGluileCys	40
Db	6011	CGGGGGCACGAGGAGCCCAAGCCGCTGCGGGCGCTGAGCGGGCAGGTGTGCGAGATATGC	6070


```
contains full-length cDNA(s): AK065475"
/codon_start=1
/product="putative RPT2"
/protein_id="BAD33627.1"
/db_xref="GI:50726106"
/translations="MKPMKGSKDPADFQSDGADVVRVVISDLATDVHVSEVKFYLIHK
FPLLSKSLIKVATEEGTDEVHIDPGPGVATFEICAKFCYGMVTLSPHNVA
ARCAAEYLENTEDVKNLI FKIDVFINSILRSWKDSIIVLQSTKALLPWSBELKVI
GRCDIASKTSVDPANVTWSYSHSRKGMSCETBEIVESTGRTS IAPKDWVVEDUCELDV
DLKYVMAVKSGRMPELIGALKAYAVRMLPDSYDALVAEDYMRNOCLVETIIM
LLPSDKTSGSCRLKLKVALVGAGQHVKEELMRRISFQLHKASVKDLLLLPAASP
SGAHVDKLVHNVIVQFVARTAMSHNGFVEKSDDKMIELNFQESTLALGEILVDGYL
SEVADPSLSTFELATVAPPAARPHVDSILYSAYDAYLKEHPNISKADKKICGLI
DVKLSPTDASMHATQNDRLRLRVQVLPFQQLRAGSSNALATDGGGHTCAKPIMKD
QSDICERRI PRHNSLNKOATSLSAREVHRKSEHGGGRNSFKDOLGQFLQSRERR
IFDKIWSKGGSGKSGSETSGSQPPLSAKADVPKPLRLNRRYSVS"
complement(join(10055..10068,13432..13638,13766..13889))
/gene="OJ1740_D06.3"
misc_feature
complement(join(10055..10068,13432..13638,13766..13889))
/gene="OJ1740_D06.3"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
14864..15751
/gene="OJ1740_D06.4"
misc_feature
14864..15751
/gene="OJ1740_D06.4"
/notes="probably inactive due to no initiation codon in CDS
pseudogene, putative HGP repeat containing protein"
/pseudo
16561..17118
/gene="OJ1740_D06.5"
misc_feature
16561..17118
/gene="OJ1740_D06.5"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative HGP repeat containing protein"
/pseudo
complement(17924..18233)
/gene="OJ1740_D06.6"
mRNA
complement(join(17924..17955,18104..18233))
/gene="OJ1740_D06.6"
/notes="start and end point are not identified"
complement(join(17924..17955,18104..18233))
/gene="OJ1740_D06.6"
/notes="similar to Oryza sativa chromosome 1, P0480E02.21"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD33628.1"
/db_xref="GI:50726107"
/translations="MKMKTIVDKRNHRHDEDDVGDPPANSSVATNDNDKDHQHLHN
ASLGGELMK"
join(18381..18410,18539..18577,19649..19732)
/gene="OJ1740_D06.7"
misc_feature
join(18381..18410,18539..18577,19649..19732)
/gene="OJ1740_D06.7"
/notes="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
25476..26366
/gene="OJ1740_D06.8"
misc_feature
25476..26366
/gene="OJ1740_D06.8"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, TNP2-like transposon protein"
/pseudo
26747..28873
/gene="OJ1740_D06.9"
mRNA
<26747..>28873
/gene="OJ1740_D06.9"
/notes="supported by full-length cDNA(s): AK065973"
27353..28669
/gene="OJ1740_D06.9"
misc_feature
/notes="putative Rim2 protein"

29731..32622
/gene="OJ1740_D06.10"
misc_feature
29731..32622
/gene="OJ1740_D06.10"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative hydroxyproline-rich glycoprotein"
/pseudo
join(36569..36654,36950..37116,37456..37510,38073..38244,
38373..38573)
/gene="OJ1740_D06.11"
misc_feature
join(36569..36654,36950..37116,37456..37510,38073..38244,
38373..38573)
/gene="OJ1740_D06.11"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(41633..42004)
/gene="OJ1740_D06.12"
mRNA
complement(<41633..>42004)
/gene="OJ1740_D06.12"
/notes="start and end point are not identified"
complement(41633..42004)
/gene="OJ1740_D06.12"
CDS
complement(41633..42004)
/notes="predicted by GlimmerM etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD33629.1"
/db_xref="GI:50726108"
/translations="MTDENPRVRAELAVTBGRGPHVAVVSLYSGNHWASREIDDAP
LUTLFFGLSLSCSGSREAREATRDDDDHQQCCSFVRVAGARSVIRLSQLASRPLFH
AKSSPNMSREMRVTARTEYHY"
42745..42927
/gene="OJ1740_D06.13"

Alignment Scores:
Pred. No.: 0 Length: 187410
Score: 4947.00 Matches: 988
Percent Similarity: 74.18% Conservative: 32
Best Local Similarity: 71.85% Mismatches: 31
Query Match: 87.59% Indels: 328
DB: 8 Gaps: 13

US-10-627-132-30 (1-1052) x AP005579 (1-187410)
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 120738 ATGAGCGGAGCGCCGGGCTGGTGGCGGTCACACACCGGACGAGCTGGTGTATC 120797
Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 120798 CCGGGGGCACGAGGAGCCCAAGCCGCTCGGGCGCTGAGCGGGCAGGTGTGCGAGATATGC 120857
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 120858 GCGGACGAGTCCGCGCCGACCCGTCGACGCGGACCTCTTCGTCGCTGCAACAGTGGCGC 120917
Qy 61 PheProValCysArgProCysTyrGluTyrGluArgGluGlyThrGlnAsnCysPro 80
Db 120918 TTCGGGTGTGCGCGCCCTGCTACGAGTACGAGCGCGCGGAGGCGCCAGACTGCCCC 120977
Qy 81 GlnCysLysThrArgTyrLysArgLeuLys----- 90
Db 120978 CAGTGAAGACCCGCTACAAGCGCTCAAGGCTAAAAAACAACACTCACATCAGCTACG 121037
Qy 90 ----- 90
Db 121038 CCCTTGATACCGGATCGCGAGGTTTCGGTTGATTGATCTTAATGCGGATGGTGGTGGT 121097
Qy 91 -----GlySerProArgValAlaGlyAspAspGluGluAspIleAsp 105
Db 121098 GGTGGTGTGTTACAGGAGCCCGAGGGTCCCGGGGACGAGGACGAGGAGGACATTGAC 121157
Qy 106 AspLeuGluHisGluPheAsnIleAspAspGluAsnGlnGln-----ArgGlnLeuGlu 123
```


Db 121158 |||||GACCTGGACGAGGTTCAACATCGACGACGAGACGAGCAGCTGCGCAGGATCAG 121217
Qy 124 GlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGly 143
Db 121218 GATGGCATGCAGACAGCCACATCCAGGCGGATGCTGCACGCGAAGATGAGTACGGG 121277
Qy 144 ArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIleIleThrGly 163
Db 121278 AGGGGCCCCGACGACGGCGACGGCAACAGCACCCCG--CTCCCGCCGATCATCACCGCG 121334
Qy 164 SerArgSerValPro----- 168
Db 121335 GCTCGCTCGT--CCCGGTACACACACACACCTCACCCCAATTCCTCTCC 121393
Qy 169 -----ValSerGly 171
Db 121394 TCTTCTCACTCAACATGTTACAGAAATGTTCTTTGTAATAAAAAAATTCAGGTGAGCGG 121453
Qy 172 GluPheProIleThrAsnGlyTyrGlyHisGlyGluValSerSerSerLeuHisIleValArg 191
Db 121454 GAGTTCGGGATATCAACAGACCATGGCCATGGCGAGTTCTCTCTTCCCTGCACAAAGCGC 121513
Qy 192 IleHisProTyrProValSerGlu----- 199
Db 121514 ATCCACCCCTACCCGGTGTCTGAGCCAGGTAGTATAACGAATTCACACAAATAATC 121573
Qy 200 -----ProGln 201
Db 121574 AATGTTCTTGTAACACACACATTTGATCATCTCAATCATCTACCGATGAATTCGTACCCAG 121633
Qy 201 ySerAlaIysTrpAspGluIysGluValSerTrpIysGluIleValSerSerSerLeuHisIleValArg 221
Db 121634 GAGTCAAAAGTGGGACGAGAGAAAGAGTGTGAGCTGGAAGAGGAGAGATGAGCACTGGAA 121693
Qy 221 sSerIysGlnGlyIleLeuGlyGlyVala---AspProGluAspMetAspAlaAspVa 240
Db 121694 ATCCAAGAGGGCATCTGTCGGCGCGCGCCCGATCCCGACGACTACGACGCGCGAGT 121753
Qy 240 1----- 240
Db 121754 CCCACTGTACGCAATCTCAGCTGAGCAGCTTACACTGATCATGATGACAACTAGTATA 121813
Qy 241 -----AlaLeuAsnAspGluAlaArgGlnProLeuSe 251
Db 121814 TTGATCATGCTGAACCTCTCTGTGGCTTGCAGG--AACGACGAGGCGAGCAGCCGCTGTC 121872
Qy 251 rArgIysValSerIleAlaSerSerIysValAsnProTyrArgMetValIleValValAr 271
Db 121873 GAGGAAGGTGTGATCGCGTCGAGCAAGGTGAACCCGTACCCGGATGGTGATCATCTCTCG 121932
Qy 271 GluValValLeuAlaPhePheLeuArgTyrArgIleLeuHisProValProAspAlaIle 291
Db 121933 TCTCTGTGTGTCTGGCTCTTCTCTCGGTACCGCATCTCCACCCCGGTGCCGACGCCAT 121992
Qy 291 eGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAs 311
Db 121993 CCCGCTGTGGCTCACCTCCATCATCTCGAGATCTGTTGCGCGGTGTCTGGATCTCTCGA 122052
Qy 311 pGlnPheProIysTrpPheProIleAspArgGluThrTyrLeuAspArgLeuSerLeuAr 331
Db 122053 CCAGTTCGCCAAGTGGTACCCGATCGACGAGGAGACCTTACCTCGACCGCGCTCTCCCTCG 122112
Qy 331 gTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrVa 351
Db 122113 CTACGAGCGGAGGGGAGCCGCTGCTGTCTGCGGGGTGGACCTGTTCTGTCAGCACCGT 122172
Qy 351 lAspProLeuIysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaVa 371
Db 122173 GGATCCGCTCAAGGAGCGCGCTGTGTACGCGCCCAACACCGTGTCTCCATCTCTCGCGT 122232
Qy 371 lAspTyrProValAspIysValSerCysTyrValSerAspAspGlyAlaSerMetLeuTh 391

Db 122233 CGACTACCCCGTCGCAAGGTGTCTCTGTACTGCTCCGACGACGCGCGCTCATGCTCAC 122292
Qy 391 rPheGluSerLeuSerGluThrAlaGluPheAlaArgIysTrpValProPheCysGlyIle 411
Db 122293 GTTCGAGTCGCTGTTCGGAGACGGCGGAGTTCGCCCGCAAGTGGTGGTCTGCTCAAGAA 122352
Qy 411 sPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuIysValAspTyrLeuIy 431
Db 122353 GTTCAGCATCAGACCGCCCGCGGAGTTCCTACTTCTCCCAAGAGGTGCGATCCTCAA 122412
Qy 431 sAspIysValGlnProThrPheValGlnGluArgAlaMetIys----- 446
Db 122413 GGACAAAGTCCATCCCAACTTCGTCAGAGCGCGCGCCATGAA--GGTAATATCGATCG 122471
Qy 446 ----- 446
Db 122472 CCATCGTATTGCTGATTCTCTAGGAGCTTGACGAAGAGCTAACTGTTGTGGGGCAATG 122531
Qy 447 -----ArgGluTyrGluGluPheIysValArgIleAsnAlaLeuValAl 461
Db 122532 GCATTTGATCGACGACGAGAGTACGAGAGTTCAGGTGAGGATCAACGCGCTGGTGGC 122591
Qy 461 aIysAlaMetIysValProAlaGluGlyTrpIleMetIysAspGlyThrProTrpProGln 481
Db 122592 GAAGCGCAGAGGTGCGCGGGAAGGTGGATCATGAAGACGGGACGCCATGCCCGGG 122651
Qy 481 yAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyIle 501
Db 122652 GAACAACACCCGCGACACCCCGGCATGATCAGGTGTTCTCGGCCACAGCGCGGCCA 122711
Qy 501 sAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluIysArgProGln 521
Db 122712 CGACACGAGGCGCAACGAGCTCCCGCGCTCTCTACGTCTCCCGTGAGAACGCGCCCG 122771
Qy 521 yPheGlnHisIleIysIysValaGlyAlaMetAsnAlaLeu----- 534
Db 122772 CTTCCAGCACCCACAAGAGGCGCGGCCCATGAACCCCTCTGTACGCCACGCCACACATC 122831
Qy 534 ----- 534
Db 122832 CTTTCATCTTCAAAACAACACAACTCTGCTCGATTGTGATAATTTCTGTTGTTGCTGGC 122891
Qy 535 --IleArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspH 554
Db 122892 AGATTGCTGTGTGCGCGCTGCTGACGAACGCGCGCTTCATGCTCAACTTGGATTGCGATC 122951
Qy 554 lserTyrIleAsnAsnSerIysAlaIleArgGluAlaMetCysPheLeuMetAspProGlnV 574
Db 122952 ACTATCTCAACACACGACGAGCCCATCAGGAGGCGATGTCTTCTCATGGATCCGAG 123011
Qy 574 alGlyArgIysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisA 594
Db 123012 TCGAGCGGAAGGTTTGTCTACGTGCGATTTCCGCGAGGTTTCGACGCGATCGACGTC 123071
Qy 594 sArgTyrAlaAsnArgAsnThrValPhePheAspIle----- 606
Db 123072 ACCGATACGCCAACCGCAACACCGCTTCTTTCGACGT--GAGCTCTCTCCCCACAACTA 123130
Qy 606 ----- 606
Db 123131 GATGAATATATACAAATCTTGAATAAATCCAGAGCAAAATCAGATCGATCGACAAATGTGAC 123190
Qy 607 -----AsnMetIysGlyLeuAspG 613
Db 123191 TGAAATTTTGTGTGTGGTCTTCTGGTGAGATCATCAGATCAACATGAAGGGCTTGATG 123250
Qy 613 lYlleGlnGlyProValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrG 633
Db 123251 GGATCCAGGGCGCGGTGTACGTTCGGGACAGGGTGGTGTTCAGGCGGCGAGCGCTGTAGC 123310
Qy 633 lYtyrAsnProProIysGlyProIysArgProIysMetValThrCysAspCysCysProC 653
Db 123311 GATCAACCCCAAGGAGCCCAAGAGGCGCCCAAGATGGTGACCTGTGCTGCTGCCCTT 123370

```
Qy 653 ysPheGlyArGlyLysArgLyshisAlaLysAspGlyLeuProGluGlyThrAlaAsp- 672
Db 123371 GCTTCGGGAGGAAGACGGAAGACGCAAGGACGGCTCCCGAGGCGCG 123430
Qy 672 ----- 672
Db 123431 ACGGGGTGAGCTCCAAATTGAGACTAAGAGAAATTAATGTTGTGAGAGATTTTCAG 123490
Qy 673 -----Met-GlyValAspSerAspLysGluMet 681
Db 123491 CTGATGGATTAATGTAAGTTTGTGTGTGGTGGATGAGGATGGAAGGACACAGAGATG 123550
Qy 682 LeuMetSerHisMetAsnPheGlyLysArgPheGlyGlnSerAlaAlaPheValThrSer 701
Db 123551 CTATGTCGAGATGAACCTTCGAGAAAGCGGTCGGGCGAGTCCGGCGGCTTCGTCAGGTCG 123610
Qy 702 ThrLeuMetGluGluGlyGlyValProProSerSerProAlaAlaLeuLeuLysGlu 721
Db 123611 ACGCTGATGGAGGAAGGCGGCTCCCGCGTCCGTCAGCCCGCGCGCTCCTCAAGGAG 123670
Qy 722 AlaIleHisValIleSerCysGlyTyrgluAspLysThrAspTrrpGlyLeuGlu----- 739
Db 123671 GCCATCCATGTCATCAGCTGGGCTACGAGGACAGACCGACTGGGGTCTCGAGGTAACC 123730
Qy 739 ----- 739
Db 123731 ACCGTTATCAGACACTAAGCCATATTACCAITGAGTGAGTTTGTGGTGAACGCCATGG 123790
Qy 740 -----LeuGlyTrrpIleTyrglySerIleThrGluAspIleLeuThrGly 754
Db 123791 ATGTGTGTGATGCGAGCTGGGGTGGATCTACCGGTGATCGATCAAGGAGGACATCTCAAGCGG 123850
Qy 755 PheLysMetHisCysArgGlyTrrpArgSerValTyrgluMetProLysArgAlaAlaPhe 774
Db 123851 TTCAGATGCACTGCGCGGGTGGAGGTCGGTGTACTGCATGCCGAGAGGCGCGGCTTC 123910
Qy 775 LysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrrpAlaLeu 794
Db 123911 AAGGGGTGAGCGCGGATCAACCTATCTCAAGCTCTCAACCAAGGTGCTCCGGTGGCGGCTC 123970
Qy 795 GlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrglyTyrglyAsnGly 814
Db 123971 GGCTCCGTCGAGATCTCTTCAGCGCGCACCGCGCTCTCAAGGCTACAGAAGCGG 124030
Qy 815 AsnLeuLysTrrpLeuGluArgPheAlaTyrglySerIleThrIleTyrglyProPheThrSer 834
Db 124031 AACCTCAAGTGGCTCGAGCGTCTCTCTACATCAACACCAACATCTACCCCTTCACTTCT 124090
Qy 835 LeuProLeuLeuAlaTyrglyThrLeuProAlaValCysLeuLeuThrGlyLysPheIle 854
Db 124091 CTCGCCCTCTCGGCTACTGCACCTACCGCGCTCGCTCTCCCTCCTCAGCGGCAAGTTCATC 124150
Qy 855 MetPro----- 856
Db 124151 ATGCCCTCCGGTCAGTCCCATCCCATCCCTGATCAATGTCCTCTTCTCCATGGATTT 124210
Qy 857 -----Ser-IleSerTh 860
Db 124211 TCCCGCCAAACTGAGTTTCAATGTTCTGAACCTTTTCTGGTGATGCAAGTATGAC 124270
Qy 860 rPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGlu 880
Db 124271 GTTTGCGAGTTTGTCTTCTCATCGGCTCTTCATCTCCATCTTCGACGCGGATCCTGGA 124330
Qy 880 uMetArgTrrpSerGlyValSerIleGluGluTrrpTrrpArgAsnGluGlnPheTrrpValI 900
Db 124331 GATGAGTGGAGCGGGGTGAGCATCGAGAGTGGTGGAGGAACGAGCAAGTTCGGGTTCAT 124390
Qy 900 eGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysValLeuAlaGlu 920
Db 124391 CGGCGGCTGTTCGGGCGCACTTGTTCGGGCTGGTGCAGGGGCTGTCTCAAGGTGCTGGCGG 124450
```

```
Qy 920 yIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAspGluPheAl 940
Db 124451 GATCGACACCAACTTCCCGTCACGTCACCAAGGCCACCGAGACGAGGACGAGATTGCG 124510
Qy 940 aGluLeuTyrrpAlaPheLysTrrpThrThrLeuLeuLeuProProThrThrLeuLeuIle 960
Db 124511 GGAGCTCTAGCCCTTCAAGTGGAGCACCCCTCTCATCCGCCACACGCTGCTCATCT 124570
Qy 960 eAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyrglnSerTr 980
Db 124571 CAACATCATCGGCTGCTCGCGGCGTCTCCGACGCCATCAACAACGCTCCGAGGCGTG 124630
Qy 980 pGlyProLeuPheGlyLysLeuPheAlaPheTrrpValIleValHisLeuTyrrpProPh 1000
Db 124631 GGGCGGCTCTTCGGGAAGCTCTTCTTCGCTTCTGGGTCTATCTCCACCTTACCCCTT 124690
Qy 1000 eLeuLysGlyLeuMetGlyArgGlnAsnArgTrrpProThrValValValIleTrrpSer 1020
Db 124691 CCTCAAGGGGCTCATGGGGAGGAGAACCCGACCCACCATTTGTCATCTGCTCGTCT 124750
Qy 1020 eLeuLeuAlaSerIlePheSerLeuLeuTrrpValArgIleAspProPheIleValArgTh 1040
Db 124751 GCTGCTCGCTCCATCTTTTCTTCTGCTCTGGGTCAAGATTGATCCCTTCCACATCAAG 124810
Qy 1040 rLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 124811 CAGGGGCTCGACGTCAGGCGAGTGGCGCATCAACTGC 124847
RESULT 5
AY095297 3277 bp mRNA linear PLN 07-OCT-2002
LOCUS Populus tremuloides cellulose synthase (Cesa2) mRNA, complete cds.
DEFINITION AY095297
ACCESSION AY095297
VERSION AY095297.1 GI:23534478
KEYWORDS Populus tremuloides (quaking aspen)
SOURCE Populus tremuloides
ORGANISM Populus tremuloides
REFERENCE 1 (bases 1 to 3277)
AUTHORS Samaga,A. and Joshi,C.P.
TITLE A new cellulose synthase gene (PtrCesa2) from aspen xylem is orthologous to Arabidopsis AtCesa7 (irx3) gene associated with secondary cell wall synthesis
JOURNAL Gene 296 (1-2), 37-44 (2002)
REFERENCE 2 (bases 1 to 3277)
AUTHORS Samaga,A. and Joshi,C.P.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2002) Forestry, Michigan Tech University, 1400
Townsend Drive, Houghton, MI 49931, USA
FEATURES
Location/Qualifiers
1..3277
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/tissue_type="xylem"
1..3277
/gene="Cesa2"
21..3119
/gene="Cesa2"
/codon_start=1
/product="cellulose synthase"
/protein_id="AAM26299.1"
/db_xref="GI:23534479"
/translation="MEASGLVAGSHNRNLVVIHGHHEHKPLKNLQGVCEICGDEI
GLTVDFGLFVACNECPKPCYEYERREGTQNCPOCKTKYRKLGKSPRVGGDEED
DVDDIIEFTIEDQDKNKLTEAMLGKMTYGRGHDEENSQPPVITGRSRPVG
EFTIGSHGEQMLSSSLKRVHPYVPSFSGNSRMDKEGKMKMDKMGHNLGPE
QDDDAEAMLEADARQPSRKRVPIASSINPVMIVARLIILAVLRYRIILHPVHDAL
GLWLTSTVCIEWPAISWILQDPKWLPIDRETYDLRLSLRYEQEGGNMLAPVDVFS
TVDPMPKPLPTGNTLLSILAMDYPVEKISCYLSDDGASMTCTFEAMSETAFARWVP
FCKFNLEPRAPRFYFTLKVDYLDKQVPTFVKERRAMKREYEEFKVRINAIKAAQK
```



```
QY 618 ValTyrValGlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProPro 637
Db 1827 GTGTATGTGGGCACAGATCGTTCCTTCAAGAGCAGCTTTGTATGCTATGACCTCC 1886
QY 638 LysGlyProLysArgProLysMetValThrCysAspCysCysProCysPheGlyArgLys 657
Db 1887 AAGGATCCAAAGGCGCCCAAGATGGAACCTGGGACTGCTGCCCATGTTTGGACGTCG 1946
QY 658 LysArgLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMetGlyValAspSer 677
Db 1947 AAAAAGAAAGATGCTAAGATGCTGGTGCAGTTGTAGAGGAATGGAT-----ATAAT 1997
QY 678 AspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAa 697
Db 1998 GACAGAGGCTGTGTATGCTCCACATGAAATTTGAAAGAGATTTGGACAAATCAGCAAT 2057
QY 698 PheValThrSerThrLeuMetGluGluGlyGlyValProProSerSerProAlaAa 717
Db 2058 TTGTAATCTTCAACTTTAATGGAAGAGGTGGTGTACCTCTTCTCGAGTCGGGAGCT 2117
QY 718 LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTropGly 737
Db 2118 CTGCTAAGAGAGCCATCCATGTATGATCAGTTGTGGATATGAAGACAAACTGAATGGGA 2177
QY 738 LeuGluLeuGlyTropIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMet 757
Db 2178 CTGAGCTGGCTGGATTTACGTTTCGATCAGCGGAGTATTTCTGACAGGTTTAAAGATG 2237
QY 758 HisCysArgGlyTropArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777
Db 2238 CATTTGCTGGCTGGAGTCTATTTACTGTATGTCAAAGAGAGCTGCATTTAAAGGTTCA 2297
QY 778 AlaProIleAsnLeuSerAspArgCysLeuAsnGlnValLeuArgTropAlaLeuGlySerVal 797
Db 2298 GCTCCCATCAATCTATCAGATCGCTTAAACCAAGTCTCGCTGGGCTCTTGGATCTGTT 2357
QY 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLys 817
Db 2358 GAAATTTTTTCAGTGTGTCACAGCCCTAATTTGGTATGGCTACAGAAGGAAGGACCTCAAG 2417
QY 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837
Db 2418 TGGCTCGAGAGTTTGGGTATGTGAACACAACTATCTACCCCTTCACCTCTCTTAGCACTC 2477
QY 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer 857
Db 2478 GTTGCACTACTGTTCCTCCCTGCGCATCTGCTGCTTACTGATAAATTTATCATGCCCCGAG 2537
QY 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877
Db 2538 ATAGCACTTTGGCAGTCTTTCTTTCATTCCTTGTGTTTGTTCATCTTTTCACCGGCG 2597
QY 878 IleLeuGluMetArgTropSerGlyValSerIleGluLufTropTrpArgAsnGluInPhe 897
Db 2598 ATTCTTGAGCTCAGATGAGCGGAGTAAGCATTCAGGAAATGGTGGAGAAACGAGCAATTC 2657
QY 898 TrpValIleGlyValSerAlaHisLeuPheAlaValGlnGlyLeuLeuLysVal 917
Db 2658 TGGGTATAGTGTGTGTCTGCTCAGCTCTTTCTGCTGTGTCAGGAGTCTTCTGAAAGTT 2717
QY 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAsp 937
Db 2718 TTAGCAGGTATCGACCTTAACTTCTGCTCATCTCAAGGCTACA-----GACGATGAC 2771
QY 938 GluPheAlaGluLeuTyrAlaPheLysTropThrThrLeuLeuIleProProThrThrLeu 957
Db 2772 GATTTTGGAGAGCTTTATGCTCTTTAAATGGCAACCCCTGCTTATCTCCCTCCCAACCACTATC 2831
QY 958 LeuIleIleAsnValIleGlyValAlaGlyIleSerAspAlaIleAsnAsnGlyTyr 977
Db 2832 TTAATCATCAACCTTTGTTGGATGTTGCTGGAGTCTCAGATGCCATAAACAATGGGTAC 2891
```

```
QY 978 GlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisLeu 997
Db 2892 CAGTCATGGGAGCCTCTATTTCGGAGAGCTCTCTTTGCTTCTGGGTGATTTGCCATCTC 2951
QY 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValIle 1017
Db 2952 TACCCCATCTCTCAAGAGTCTAATGGGAGGCAAAACAGGACACCGACTATTGTGGTTATA 3011
QY 1018 TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTyrValArgIleAspPropheIle 1037
Db 3012 TGGTCAGTGTCTGCTGCTTCCATCTTCTCTTGTGGTCCGGATTGATCCATTGTG 3071
QY 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 3072 ATGAAACACAGGAGCCTGCACACCAAGCAATGTGGACTCAACTGT 3116

RESULT 6
BT013780 3414 bp mRNA linear PLN 11-MAY-2004
LOCUS Lycopersicon esculentum clone 132670R, mRNA sequence.
DEFINITION BT013780
ACCESSION BT013780
VERSION BT013780.1 GI:47105195
KEYWORDS FLI_CDNA.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
REFERENCE 1 (bases 1 to 3414)
AUTHORS Kirkness, E.F., Wang, W. and Vazeille, A.
TITLE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
SUBMITTED (11-MAY-2004) The Institute for Genomic Research, 9712
MEDICAL CENTER DRIVE, ROCKVILLE, MD 20850, USA
FEATURES
    Location/Qualifiers
        1..3414
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /db_xref="taxon:4081"
            /clone="132670R"
            /tissue_type="fruit"
            /note="TZXB89"

ALIGNMENT SCORES:
Pred. No.: 0 Length: 3414
Score: 4784.50 Matches: 880
Percent Similarity: 90.72% Conservative: 78
Best Local Similarity: 83.33% Mismatches: 19
Query Match: 84.71% Indels: 19
DB: 8 Gaps: 7

US-10-627-132-30 (1-1052) x BT013780 (1-3414)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 66 ATGGAAGCTAGTCAGGCTCTAGTTGCTGGTTCTCATTAACCGGAATGAACCTTGTGCTATT 125
QY 21 ArcGlyHisGluAspProLysProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 126 CATGGCCATGAAGAGCATTAAGCCATTGAAGGATTTGAGTGGACAGATTTTGTGACATATGT 185
QY 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 186 GGGGATGAATAGGCTTTAACAGTGGATGGTGTCTTTTGTGGCTTCGAATGAGTGTGGT 245
QY 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
Db 246 TTCCCTGTGTGTCGGCTCTGCTACGAGTATGAGAGAAAGGAAGAACTCAACAGTGTCCA 305
QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 306 CAGTGCAGACCAAGATCAAAACGCTCTTAAGGGAGTCCCAAGGGTGGCGGGAGATGATGAT 365
```

Qy	101	GluGluAspIleAspAspLeuGluHuHisGluPheAsnIleAspAspGluAenGlnArg	120
Db	366	GAGCAAGATATTGACGATATTGAACTGAATTCAAAGTTGATGATGACCAAAAC-----	419
Qy	121	GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet	140
Db	420	-----AGAAATAGAAACAAATTGTCGAGACCAATTCCTCCATGGTAAAGATG	461
Qy	141	SerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIle	160
Db	462	ACATATGGAAGAGGACCAAGATGAAGAT-----TCCGCTCAGTACCTCCCTGCTC	512
Qy	161	IleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGly	180
Db	513	ATTGCTGGAAACCCGCTCACATCCGGTTAGCGGTGAATTCCTCAATATCGAAAC--CATGGA	569
Qy	181	HisGlyGlu-----ValSerSerSerLeuHisGlyArgIleHisProTyrProValSer	198
Db	570	AATGGAGAGCAGATGTTGGGATCTTCGCTTCACAAAGGATACATCCCTATCCAGCTTCT	639
Qy	199	GluProGlySerAlaLysTrpAspGluLysGlyValSerTyrTrpLysGluArgMetAsp	218
Db	630	GAATCTGGAAAGTCAAGATGGATGATAGAAAGGGGAGGGTGGAAAGAAAGATGAA	689
Qy	219	AspTrpLysSerLysGlnGlyIleLeuGlyGlyGlyAlaAspProGluAspMetAspAla	238
Db	690	GATTGGAAATTTTCAGCAAGGACACCGCGGGCAAGATTATGAT---GACTCTGCTGACGTT	746
Qy	239	AspValAlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSer	258
Db	747	GATATGTCCATGGTGGATGAAGCAAGACAGCCTTTATCAAGAAAGTGCCCAATAGCATCA	806
Qy	259	SerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePhe	278
Db	807	AGCAAGATCAATCTTACCGGATGGTGAATGTAGCTAGGCTGGTGATCTGGCGGCTTC	866
Qy	279	LeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerIle	298
Db	867	CTTCGTTATAGAATTTCTGAATCCAGTGCATGCGATTCGCTCTCTGGTTAACTTCTATC	926
Qy	299	IleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePro	318
Db	927	ATCTGTGAGATCTGGTTGGCTCTCTTGGAATCCTTGATCAGTTTCCCAATGGTTCCT	986
Qy	319	IleAspArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluPro	338
Db	987	ATCGACGCTGAGACTTACCTTGATCTCTCTCTCAGGTATGAGAGGGAGGAGAACCA	1046
Qy	339	SerLeuLeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProPro	358
Db	1047	AATATGTAGACCCGGTAGATGTCTTTGTGAGTACAGTGGATCCTATGAAGGAACCTCCT	1106
Qy	359	LeuValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysVal	378
Db	1107	CTTGTTACAGCTAATACAACTCCTCAATATTAGCAATGATTAACCTGTTGACAAAT	1166
Qy	379	SerCysTyrValSerAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThr	398
Db	1167	TCATGTACCTATCTGATGATGAGCTTCAATGTGCACCTTTGAGACACTATCAGAACT	1226
Qy	399	AlaGluPheAlaArgLysTrpValProPheCysLysLysPheGlyIleGluProArgAla	418
Db	1227	GCAGAAATTTGCTAGAAAATGGTTCCCTCTCTGCAAGAAGTTTGCTATAGAGCCTCGAGCT	1286
Qy	419	ProGluPheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPhe	438
Db	1287	CCAGAAATTTTACTTCTCCCTTAAGATTGATTAATCTCAAGACAAAAATCCAGCAACATTC	1346
Qy	439	ValGlnGluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAla	458
Db	1347	GTTAAGAGGAGACGAGGCATGAGAGAGAGTACGAGAGNATTCAGGTTAGATCAATGCA	1406
Qy	459	LeuValAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrPro	478

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 1 ATGGAGCTAGCGCGGCTTTGTTCGCCGGTTCTCATAAACCGGTAACTAGTCGTCATT 60
Qy 21 ArgGlyHisGluAspProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 61 CACAACCATGAAGCCAAAGCCCTCTGAAGAACTCTAGATGGACAATCTCTGTGAGATATGT 120
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 121 GGAGATCAGATTGGTTTAAACAGTAGAAGGAGACCTCTTCGTAGCTTGGCAATGAGTGTGGT 180
Qy 61 PheProValCysArgProCysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysPro 80
Db 181 TTTCCGGCGGTGTAGACCTTGTCTATGAGTACGAGAGAGAGAGAGGAAACCAAACTGTCTCT 240
Qy 81 GlnCysIleThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
Db 241 CAGTGTAGACTCGTTACAGCGTCTCAGAGGAGCCCAAGAGTGGAGGGAGATGAAGAC 300
Qy 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db 301 GAAGAAGATATTGATGATATTGAGTATGAATTTAATATCGAACATGAACAAGAT----- 354
Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 355 -----AAGCATAAAGCAATTTCTGCTGAGGCTATGCTTTTATGGGAAATG 396
Qy 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProfile 160
Db 397 AGCTATGGAGAGGTCCTGAGATGATGAGATGGG-----AGATTCCCACCTGTT 447
Qy 161 IleThrGlySerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGly 180
Db 448 ATAGCTGTGTGTCAT-----AGTGAGAGAAATTTCCAGTTGGAGGAGTTATGTT 495
Qy 181 HisGlyGluValSerSerSerLeuHisLysArgIleHisProTyrProValSerGluPro 200
Db 496 AATGGAGAA-----CATGGGCTTCATAAAGCGTGTGACCCCATATCCATCATCTCTGAAGCT 549
Qy 201 GlySerAlaLysTyrAspGluLysGluValSerTyrLysGluArgMetAspAspTyr 220
Db 550 GGGAGT-----GAGGAGAGATGCGCGGAGAAAGAAAGTGGATGACTGG 588
Qy 221 LysSerLysGlnGlyIleLeuGlyGlyAlaAspProGluAspMetAspAlaAspVal 240
Db 589 AAGCTCCAGCATGGAAATCTT-----GGGCCAGAACACAGATGAT--GATCCTGAGATG 639
Qy 241 AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlaSerSerLys 260
Db 640 GGACTGATCGACGAGGCGCGCAGCCACTCTCGCGGAAAGTTCCCATTTGCCCTCAAGCAAG 699
Qy 261 ValAsnProTyrArgMetValIleValArgLeuValValLeuAlaPhePheLeuArg 280
Db 700 ATCAATCCATATCGGATGGTCACTGCTGCTAGGCTTGATCTTAGCAGTTTTCCTGCGG 759
Qy 281 TyrArgIleLeuHisProValProAspAlaIleGlyLeuTyrLeuValSerIleIleCys 300
Db 760 TATAGGCTCTTGAATCCAGTCAGTATGCTCTGGGATTTAGCTGACCTCTGTGATCTGT 819
Qy 301 GluIleTyrPheAlaIleSerTyrIleLeuAspGlnPheProLysTyrPheProIleAsp 320
Db 820 GAAATCTGTTTCGCTGTCTCTTGGATTCTTGATCAGTTCCCGCAAGTGTGTCCTATTGAA 879
Qy 321 ArgGluThrTyrLeuAspArgLeuSerLeuArgTyrGluArgGluGlyGluProSerLeu 340
Db 880 CGTGAGACTATCTAGATCGGCTTTCCTCTAGGTACGAGAGAGAGGTAACCAAAATATG 939
Qy 341 LeuSerAlaValAspLeuPheValSerThrValAspProLeuLysGluProProLeuVal 360
Db 940 CTTCGCCCTGTAGATGCTTTGTCTAGTACGCTGGACCCATTTGAAGGAGCCTCCCTCGTC 999
Qy 361 ThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCys 380

Db 1000 ACATCCCAACACTGTGCTGTCAATCTTGGCCATGACTACCCAGTTGAGAAAATCTCTCCG 1059
Qy 381 TyrValSerAspAspGlyAlaSerMetLeuThrPheGluSerLeuSerGluThrAlaGlu 400
Db 1060 TATGCTCTCTGACGAGCGTGTCAATGCTTACATTGGAATCTCTCTCGAANAACCTGCTGAG 1119
Qy 401 PheAlaArgLysTyrProValProPheCysLysLysPheGlyIleGluProArgAlaProGlu 420
Db 1120 TTTGCAAGAAAATGGGTTCCCTCTGTAAGAAAATTTCTCATAGAGCCACCGGCCACCGGAG 1179
Qy 421 PheTyrPheSerLeuLysValAspTyrLeuLysAspLysValGlnProThrPheValGln 440
Db 1180 ATGTACTTTCAGTTGAAAGTTGATTATCTCCAGACAAAGTCCACCCCAACATTTGTTAAG 1239
Qy 441 GluArgArgAlaMetLysArgGluTyrGluGluPheLysValArgIleAsnAlaLeuVal 460
Db 1240 GAACGTCGAGCCATGAAGAGAGATATGAGGAGTTCAAGGTCAGAGTCAACGCTCAAGTG 1299
Qy 461 AlaLysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThrProTyrPro 480
Db 1300 GCGAAGGCTCAAAAGGTTCTCTAGAAAGTTGGATCATGCAAGATGGAACACCGGTGGCCA 1359
Qy 481 GlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGly 500
Db 1360 GGGAAACCAACCAAGGACCCCGGTATGATCCAAGTCTTCTCGGCCACAGCGGAGGA 1419
Qy 501 HisAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgPro 520
Db 1420 TTTGATGTCGAGGGCATGAGCTTCTCGGCTGTGTACGTGTCCTCGTGAAGCGTCTCT 1479
Qy 521 GlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuIleArgValSerAlaVal 540
Db 1480 GGTTTTCAACACCAAGGTCGGCCATGAATGCCCTGTTCTGAGTGGCAGCGCGTA 1539
Qy 541 LeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLys 560
Db 1540 CTCACAAATGCTCTTTCATGCTGAACTTGGACTGTGATCACTATGATATAAACACAGCAAG 1599
Qy 561 AlaIleArgGluAlaMetCysPheLeuMetAspProGlnValGlyArgLysValCysTyr 580
Db 1600 GCCGTGAGGAGAGCAATGTGTTTTCATGGATCCTCAGATTGGAAAGAGGTCGTCTAT 1659
Qy 581 ValGlnPheProGlnArgPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsn 600
Db 1660 GTTCAGTTCTCCCTCAAGGTTTGTATGTCATTCACAAACGATCGTTACGCCAACAGAAAC 1719
Qy 601 ThrValPhePheAspIleAsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrVal 620
Db 1720 ACAGTCTTCTTTGATATCAATATGAAAGGTCCTAGATGGAATCCAAGGTCCTAGTTTACGTT 1779
Qy 621 GlyThrGlyCysValPheArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyPro 640
Db 1780 GGTACTGTTGTGTTTTCAAACGACAGCTCTGTATGTTATGAACACCAAGAGGTCCT 1839
Qy 641 LysArgProLysMetValThrCysAspCysProCysPheGly-----ArgLysLys 658
Db 1840 AAACGTCCAAAGATGATGAAGCTGTGTTGTTGCTTGTGCTTGGCGCGCGGAGAAAGAAT 1899
Qy 659 ArgLysHisAlaLysAspGlyLeuProGluGlyThrAlaAspMet---GlyValAspSer 677
Db 1900 AAGAAATTTTCCAAAGAATGACATGAATGGTGACCTAGCAGCCCTTGGAGGAGCAGAGGT 1959
Qy 678 AspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyGlnSerAlaAla 697
Db 1960 GATGAGAGAGCATTTGATGTCGAAATGAACTTTGAGAAAACATTTGGCAATCATCCATC 2019
Qy 698 PheValThrSerThrLeuMetGluGluGlyValProProSerSerSerProAlaAla 717
Db 2020 TTTGTAACTCACTTTTGTATGGAAGAGGTTGTTGCTCTCGTCTCATCAAGTCTGCAAGT 2079
Qy 718 LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysThrAspTyrGly 737

Db 2080 CTCTTAAAGAGGCAATCCATGTCATTAAGCTGCGGTTATGAAGACAGACTGAGTGGGA 2139

Qy 738 LeuGluLeuGlyTrpIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLeuMet 757

Db 2140 ACTGAGCTGGTGGATCTATGGCTCTATCACAGAGGATATTTGACGGGATTCAGATG 2199

Qy 758 HisCysArgGlyTrpArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySer 777

Db 2200 CATTCGCGTGGATGGAGGCTCTATTTACTGCTGCTAAGAGCGCTGCATTCAAAGGTTCA 2259

Qy 778 AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVal 797

Db 2260 GCTCCTATTATCTATCAGACAGGTTAAACAGGTTTTGGCTTGGGCACATTGATCGGTT 2319

Qy 798 GluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817

Db 2320 GAGATATTTTACCGCGGACAGTCTCTCTGGTATGGCTACAAAGAGGCAAACTCAAG 2379

Qy 818 TrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLeu 837

Db 2380 TGGCTTGAGCGTTTTGCTTATGCGCAACACAACTATACCCCTTCACATCTATACCACCT 2439

Qy 838 LeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetProSer 857

Db 2440 CTTGCCCTACTGATCTCTCCAGCCATCTGCTCTTACTGACAAATTCATCATGCCACCG 2499

Qy 858 IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrGly 877

Db 2500 ATAAGCACATTTGCTAGTCTCTTCTTCATCTCTGTTTATGTCATCTGTTAAACGGGA 2559

Qy 878 IleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnPhe 897

Db 2560 ATCTTGCAATTGAGATGAGCGGAGTTAGCATTTAGGCAATGGTGGAACAGCAATTC 2619

Qy 898 TrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysVal 917

Db 2620 TGGGTCAATTGGAGGAATCTCAGCTCATCTCTTTCGGGTTGTCCAAAGTCTCCTCAAAAT 2679

Qy 918 LeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspAsp 937

Db 2680 TTAGCAGGCATTTGCACAACTTCACCGTTCACATCAAAAGCGNACA-----GATGATGAT 2733

Qy 938 GluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrLeu 957

Db 2734 GACTTTGGAGAACTTTACGCATTCAAATGGACAACACTGCTGATCCCTCCACCAACTGTC 2793

Qy 958 LeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyTyr 977

Db 2794 TTAATCATAAACATTTGTTGGCGTTGTTGCGAGGCATCTCAGATGCCATTAAACAATGGATAT 2853

Qy 978 GlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeu 997

Db 2854 CAGTCTTGGGAGCCTCTATTTTGGTAAACTCTTCTCTCTTTTGGGTCATGTTCATCTC 2913

Qy 998 TyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValIle 1017

Db 2914 TACCATTCCTCAAAGGCTGATGGGTAGACAGAACAGACACCAACCATTTGGTGAT 2973

Qy 1018 TrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspProPheIle 1037

Db 2974 TGGTCATGTTATTTGGCATCTATCTCTCTTTGTTGGTAAAGATTTGATCTCTTTGTG 3033

Qy 1038 ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052

Db 3034 CTCGAAGACCAAGAGCACTGACACTTCCAAAGTGTGGCATCAACTGC 3078

RESULT 8

AV139754

LOCUS AV139754 3355 bp mRNA linear PLN 02-SEP-2002

DEFINITION Arabidopsis thaliana AT5g17420/t10B6_80 mRNA, complete cds.

ACCESSION AV139754

VERSION AV139754.1 GI:22654964

KEYWORDS FLI_CDNA.

SOURCE
ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS

1 (bases 1 to 3355)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y.,
Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,
Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,
Theologis, A. and Ecker, J.R.

TITLE
JOURNAL

Arabidopsis cDNA clones

REFERENCE
AUTHORS

2 (bases 1 to 3355)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y.,
Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,
Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,
Theologis, A. and Ecker, J.R.

TITLE
JOURNAL

Direct Submission

Submitted (09-AUG-2002) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA' : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,
Cheuk, R., Kim, C.J., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M.,
Deng, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Quach, H.L.,
Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,
Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and
Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source

1. .3355

Location/Qualifiers

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="RAFL09-35-F05 (R22199)"

/ecotype="Columbia"

46. .3126

/note="cellulose synthase catalytic subunit (IRX3)"

/codon_start=1

/product="AT5g17420/t10B6_80"

/protein_id="AAM98075.1"

/db_xref="GI:22654965"

/translation="MEASAGLVASHNRNELVVIHNEBPKPLKLDGQFCEICGDOI
GLTVEGDLFVACNEGCPACPCYERREGTCNCPCKTRKRLGSPRVEGDDE
DIDDIEYENIEHQDKHSAEMLYGMSYGRGPDDENGFPVILAGHSGEPV
GGYNGEHLKRVHPYFSSAGSGEGRMRMDMKLQHLGNLGPEDDDPEMGLIDE
ARQPLSRKPIASCKINPYRMVILAVILAVFLRYRLNPDVHDLGLWTSVCEI
FAVSTLDGFPKFPFIERETYLDRLSIEREGEPNMLAPDVFTVDLKEPPLVT
SNVLSILAMDYKVEIKSCVSDGASMLTFESLSEAFARAKWPFCKKFSIEPRAP
ENYFTLKVDLQDKVHTFVKERRAMKREYBPKVINAQAKAVPGLGWTMODCT
PFPNGNTKDHGMQVFLHSGSPDVEGHELRLVYVSEKRPQGFQHHKAGAMALV
RVAGLVNTAFFMLNLOCDHYVNSKAVREAMCFMLMDPQIGKKYVQVQFPQRPIDTN
DRYANRNTVFFDINMKGLDGIQGFVYVVGTCVFKRQALYIGYEPKPKPKMISCGCC

Db	2974	TGGTCAGTGTTATTGGCATCTATCTCTCTTTGCTTTGGGTAAAGAAATGATCCTTTTGTG	303
Qy	1038	ValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys	1052
Db	3034	CTCAAGACCAAGGACCTGACACTTCCAAGTGTGGCATCAACTGC	3078
RESULT 10			
AY639654		3911 bp	linear
LOCUS	Pinus radiata cellulose synthase catalytic subunit (Cesa1) mRNA,		PLN 29-JUN-2004
DEFINITION	complete cds.		
ACCESSION	AY639654		
VERSION	AY639654.1	GI:49182339	
SOURCE	Pinus radiata (Monterey pine)		
ORGANISM	Pinus radiata		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus		
AUTHORS	1 (bases 1 to 3911)		
TITLE	Klein,A.S., Tibbitts,J., Steven,R. and Anthony,B.		
JOURNAL	Cellulose synthase genes in Conifers: what we know and what we need to learn		
REFERENCE	(in) Hayashi,T. (Ed.); PLANT CELL WALLS; (2005) In press		
AUTHORS	2 (bases 1 to 3911)		
TITLE	Klein,A.S., Tibbitts,J., Steven,R. and Anthony,B.		
JOURNAL	Direct Submission		
FEATURES	Submitted (27-MAY-2004) Biochemistry and Molecular Biology, University of New Hampshire, 46 College Road, Durham, NH 03824, USA		
source	Location/Qualifiers		
gene	1. .3911		
5'UTR	/organism="Pinus radiata"		
CDS	/mol_type="mRNA"		
	/db_xref="taxon:3347"		
	1. .3911		
	/gene="Cesa1"		
	/note="PrCesa1"		
	1. .168		
	/gene="Cesa1"		
	169. .3423		
	/EC_number="2.4.1.12"		
	/note="glycosyltransferase"		
	/codon_start=1		
	/product="cellulose synthase catalytic subunit"		
	/protein_id="AA157672.1"		
	/db_xref="GI:49182340"		
	/translation="MEASGLVAGSHNRNEFVVIHGHEPKPLNTLSGHVCQICGEDV		
	GLNTDGEILFVACNECGPVCPCYEYERREGNSQPCQNTYRIKRGKSPRGVGGDDDE		
	DVDDLEHFNFAEQORNQITTEAMLGRMSYGRGPDENSQIAHNPELPQIPVLAN		
	GHSVGYSPITSYADNOLLANPAWLKRVHSPSPSGRIIMDPNRDIGSVGFGNWSW		
	KERGQYGSKENKSQLDMTEGRYQYNGGAPNPEPIYIDPMNPTDEARQPLSRKVP		
	IPSSKINPYRVIIRLIVLGFILRYEREGFSLAPFDLVFVDPKPEPLVNTANTVLSLV		
	PWLPISRETYLDRSLRYEREGFSLAPFDLVFVDPKPEPLVNTANTVLSLV		
	DYPDVNSCYTSDGASMLTESISETSEFAKWPVFCFKFDIEBRAPEIYFSGKIDY		
	LKDEQPTFVKERRAMKREYEFKVRINLIVAKASVPEKGTWQDGTGPPNPNRDRH		
	PGMIOVLHGSGGLDTEGNEPLRVLYVSREKRGFQSHHAKAGMNLVRYSAVLNTPA		
	FMLNDDCHYINNSKAIKREGCMFMDPQVRKVCYVFPQRFQDIDRNDRYANRNTVP		
	FDINNKSGDGIQGPVYGTGCMFRQALYGGPPKPRPMWTCDCLPCCGPKKSP		
	KNNSSKAGAGIPAPYALNDGIEEGVEGYDDERALLMSQLDFEKFGSSAFVQSTLME		
	NGCVPTQANPAELLKEAHLVISCYGEYDTEWGEKELWYIGVSTEDILTGFQMKTRH		
	STYCMKPEAAFGKGAIPNLSDRLNQLRWALGSVEIFMSRHCPITWYGVGGLKWLFR		
	AYINTIVYPTSLPLIAXCTLPAYSLLTGKEVYIPISTFASLFFIALFISIPATGILE		
	MWGSVSEIIEWNRNEQFWVIGVSAHFPAVLOGLKVLKLAGIDTNTFTVTKASDGEFG		
	ELYAPFKWTKTLLIPTTLVLINLVGVVGVADAINNGFSWGLGLKUFFFAFPWLVILH		
	PFLKGLMGQRNRTTIVIVINSILLASVPSLFWVRIDPFLSKVGKGPDTKQCGINC"		
	3424. .3879		
	/gene="Cesa1"		
3'UTR			
ORIGIN			
Alignment Scores:	0	Length:	3911
Pred. No.:			

Score: 4483.00 Matches: 841
Percent Similarity: 84.56% Conservative: 90
Best Local Similarity: 76.39% Mismatches: 104
Query Match: 79.37% Indels: 66
DB: 8 Gaps: 16

US-10-627-132-30 (1-1052) x AY639654 (1-3911)

```
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeuIle 20
Db 169 ATGGAAGCAGCGCGGCTTGGTTGCGGTTCTCATACAGAAACGAGTTCGTGTCATC 228
Qy 21 ArgGlyHisGluAspProLeuArgAlaLeuSerGlyGlnValCysGluIleCys 40
Db 229 CATGACATAGAGGCGCAAGCCCTTGAACAGTGTGAGTGCCACGCTCGCCGATTTGT 288
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 289 GCGAGGAGCGTCGGCTTAACACAGACGCGGAGCTGTTGCTGTGAATGAGTGGCGG 348
Qy 61 PheProValCysArgProCysTyrrGluTyrrGluArgArgGluGlyThrGlnAsnCysPro 80
Db 349 TTTCTGTCTGTCGCGCGTGTATGATGATGACGAGACGAGAGGAATCAGTCGTGCGCG 408
Qy 81 GlnCysLeuThrArgTyrrLeuArgLeuGlySerProArgValAlaGlyAspAsp 100
Db 409 CAGTGCATATCTCGTTTCAAGCGTCAAAAAGGAGTCCACGGTGGAGGATGATGAT 468
Qy 101 GluGluAspIleAspAspLeuHisGluPheAsnIleAspAspGluAsnGlnArg 120
Db 469 GAAGAGAGCTTGATGACATAGACATGAATTTAATGTGGAGGCTCAGCAAGAAACAGG 528
Qy 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
Db 529 CAG-----CAGATCACCGAGGCGATGCTCCACGCGCGCATG 564
Qy 141 SerTyrrGlyArgGlyProAspAspGlyAspGlyAsn-----AsnThrProGlnIlePro 158
Db 565 AGCTATGGCCGAGGTCCGACGACGAAATTCGACAGATGTCATATAATCCAGAGCTTCCT 624
Qy 159 -----ProIleThrGlySerArgSerValProValSerGlyGluPheProIle 175
Db 625 CCGCAGATTCCTGTACTTGCAACCGGCCACTCGGTT---GTGAGTGGGAGATTCCA--- 678
Qy 176 ThrAsnGlyTyrrGlyHisGlyGluVal-----SerSerSerLeuHisLysArgIleHis 193
Db 679 ACGTCATACTACGACAGCAACCAATTGCTTCCCAACCTCGCAATGCTGAAGCGTGTGCAT 738
Qy 194 ProTyrrProValSerGluProGlySerAlaLys-----TrpAspGluLysLysGlu--- 210
Db 739 -----CCAGCTCCGAGCCGGGAGTGGAAAGGATCATCATGATCCAAACAGGGATATT 792
Qy 211 -----ValserTrpLysGluArgMetAspAspTrpLysSerLys 223
Db 793 GGTTCATTATGCTTTGGGAACGTGCTCTTGGAAAGGAGCGAGTGATGTTATTAATTCGAAG 852
Qy 224 GlnGlyIleLeuGly-----GlyGly--- 230
Db 853 GAAACCAAAATCAGGCCAGTGGATATGACGGAAGGAGATATCAATATAATGGGGGTTT 912
Qy 231 -----AlaAspProGluAsp---MetAspAlaAspValAlaLeuAsnAspGluAlaArg 247
Db 913 GCACCAATAGCTCGAAGATTATATTGATCCCATATGCCCAATGACCGGATGAAGCAAGG 972
Qy 248 GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrrArgMetVal 267
Db 973 CAGCCTGTCTCCGAAAGTCCCAATTCCTTCAAGCAAAATTAATTCATACCGAATGGTC 1032
Qy 268 IleValValArgLeuValValAlaPhePheLeuArgTyrrArgIleLeuHisProVal 287
Db 1033 ATTGTAATTCGACTGATAGTCTGGGTATTTTCTCCGCTATCTCTCCTGAATCCAGTG 1092
Qy 288 ProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSer 307
```

```
Db 1093 AAGAAATGCATATGGGCTCTGGGCCACTTCTATCGTTTGTGAAATCTGGTTTGCCTTGTCA 1152
Qy 308 TrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrrLeuAspArg 327
Db 1153 TGGATTCCTGATCAGTTTCCCAAGTGGTGCCTATCAGTCGTGAACGATATCTTGATCGA 1212
Qy 328 LeuSerLeuArgTyrrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe 347
Db 1213 CTGTCAATTAAGGTACGAAACGAGAGGCGAACCATCAATGCTTGACCTGTGTGACCTCTTT 1272
Qy 348 ValSerThrValAspProLeuLysGluProLeuValThrAlaAsnThrValLeuSer 367
Db 1273 GTGAGTACTGTAGATCCACTGAAGAGGCCCTCTTTGGTTTACTGCTCAATATACAGTATTATCA 1332
Qy 368 IleLeuAlaValAspTyrrProValAspLysValSerCysTyrrValSerAspAspGlyAla 387
Db 1333 ATCCCTTTCAGTACAGTACCTCTGATAGCAATGTGTCTCTGTTATGTCTCTGATGACGAGCG 1392
Qy 388 SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro 407
Db 1393 TCGATGCTTACTTTTGAATCTCTCTGAGACCTCAGAAATTTGCCAGAAAATGGGTACCA 1452
Qy 408 PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrrPheSerLeuLysVal 427
Db 1453 TTTCTGCAAGAAATTCGACATTCGACCTCGCGCTCCCGAAATCTATTTCTCTCAGAAAAT 1512
Qy 428 AspTyrrLeuLysAspLysValGlnProThrPheValGlnArgArgAlaMetLysArg 447
Db 1513 GACTATCTGAAGGACAAATTTCAACCACCTTTGTCAAGAGCGCCGGCCATCAAGAGA 1572
Qy 448 GluTyrrGluGluPheLysValArgIleAsnAlaLeuValAlaLysAlaMetLysValPro 467
Db 1573 GAATATGAAGAAATTCAGGTGCGCATCAATCGGTGTTGTTGCAAAAGGCTCTAAAGTGCC 1632
Qy 468 AlaGlyLysTrpIleMetLysAspGlyThrProTyrrProGlyAsnAsnThrArgAspHis 487
Db 1633 AAGGAAGATGGACAAATGCAAGACGGTACGCTTGGCTGGTAAATAATACCCGTGACCAT 1692
Qy 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGlu 507
Db 1693 CCTGGTATGATCCAGTGTCTTGGGTACAGTGGCGGCTCGCATACAGAAAGGCAATGAG 1752
Qy 508 LeuProArgLeuValTyrrValSerArgGluLysArgProGlyPheGlnHisLysLys 527
Db 1753 CTTCCTCGGCTAGTATATGTTCTCGTGAAGAGACCTGGTTTCCAGCATCACAAGAAG 1812
Qy 528 AlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMet 547
Db 1813 GCCGGTGCCCAATGCTTTGGTTGCGGTTTCTGCTGTGCTCACAATGCTCCATTTATG 1872
Qy 548 LeuAsnLeuAspCysAspHisTyrrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
Db 1873 CTGAATCTGGATGTGATCACTACATTAACAACAGCAAGGCATCAGGAAGGCAATGTGC 1932
Qy 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrrValGlnPheProGlnArgPhe 587
Db 1933 TTTATGATGATCCTCAGGTGGGAGAAAAGTCTGTTATGTTCCAAATTCCTCCATCAGAGATTC 1992
Qy 588 AspGlyIleAspValHisAspArgTyrrAlaAsnArgAsnThrValPheAspIleAsn 607
Db 1993 GATGGTATTGATCGCAATGACCGTTACGCCAAATCGAAACACCGGATTTCTTTGATATCAAC 2052
Qy 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrrValGlyThrGlyCysValPheArg 627
Db 2053 ATGAAGGTCTGGATGGAAATTCAGGGGCTGTATATGTTGGGAACGTGATGATGATTCAGA 2112
Qy 628 ArgGlnAlaLeuTyrrGlyTyrrAsnProProLysGlyProLysArgProLysMetValThr 647
Db 2113 AGACAAGCTCTATATGGGTATGGGCTCCCAAGGCCCAAAACGTCCTCAAGATGGTGACC 2172
Qy 648 CysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLys----- 663
```

```
Db 2173 TGTGATTGTCTCCCTTGTGGGTCTCGTAAAGAGTCTCCGAGAAAAATAGTAGCAAG 2232
Qy 664 -----AspGlyLeuProGluGlyThrAla 671
Db 2233 AAAAGTCAGGAATCCAGCTCCCGCTCAAACTCGACCGGATCGAGGAGAGTAGAA 2292
Qy 672 AspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAenPheGluLysArg 691
Db 2293 -----GGTTATGATGACGAAGAGCATTTGATGAGCACTAGACTTCGAGAGAAG 2346
Qy 692 PheGlyClnSerAlaAlaPheValThrSerThrLeuMetGluGlyValProPro 711
Db 2347 TTGGCCAGTCTTCAGCTTTTGTTCATCCACTCTGATGGAGAAATGGTGTTCGCGCA 2406
Qy 712 SerSerProAlaAlaLeuLeuLysGluAlaLeuHisValIleSerCysGlyTyrGlu 731
Db 2407 ACAGCAATCCAGTGAATTTGGAAGGAGCTATTCTATGTCATCAGCTGTGATATGAA 2466
Qy 732 AspLysThrAspTrpGlyLeuGluLeuGlyTyrIleTyrGlySerIleThrGluAspIle 751
Db 2467 GACAAAACGGAATGGGGAAGAGCTTGGATGATCTATGGATCAGTCACAGAGGACATT 2526
Qy 752 LeuThrGlyPheLysMetHisCysArgGlyTyrArgSerValTyrCysMetProLysArg 771
Db 2527 CTGACTGGATTCAAGATGCACATCGAGGCTGGCGTCCATTTACTGTATGCCCAACGA 2586
Qy 772 AlaAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg 791
Db 2587 GCAGCATTCAGAGGCTCTCCATCAATCTATCAGACCGTTTGAACCAAGGTGTGCGT 2646
Qy 792 TrpAlaLeuGlySerValGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyr 811
Db 2647 TGGGCTTTGGGATCAGTAGAAATTTTCATGAGCAGACATTCGCCCAATCTGCTATGCGTAT 2706
Qy 812 LysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 831
Db 2707 ---GGGGAGGCTGAAATGGCTTGAAGATTTGGCTATATCAACCACTATGCTATCCA 2763
Qy 832 PheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGly 851
Db 2764 TTCACCTCTCTCCACTATGGCTATTGCACACTTCCAGCCGTCAGTTTGTCTCATGGC 2823
Qy 852 LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMet 871
Db 2824 AAATTTGTGATCCCTCAGATCAGTACTTTTGGCAAGTCTATTTTATAGCTCTTTTCATC 2883
Qy 872 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluTyr 891
Db 2884 TCAATTTTGGCCACTGGTATTTCTGGAAATGAGGTGGAGTGGAGTGAAGCAATGG 2943
Qy 892 TrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValVal 911
Db 2944 TGGCGAATGAAACAGTCTGGGTATTGGAGGGTTCCTGCACATTTTTCAGTTATT 3003
Qy 912 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 931
Db 3004 CAAGGTCTGCTCAAGGTACTGGCAGGCATTTGATACAAATTTTCAGTCACTGCCAAGCA 3063
Qy 932 ThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeu 951
Db 3064 TCA-----GATGACGGTGGTTGGGGAACTGTATGCATTCAAATGGACCACTCCCT 3117
Qy 952 IleProProThrThrLeuLeuIleAsnValIleGlyValValAlaGlyIleSerAsp 971
Db 3118 ATTCTCTTACACCTGCTTGTTCATCAACTGTGTGGGGTGGTGTGTGGCGTAGCAGAT 3177
Qy 972 AlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 991
Db 3178 GCAATCAACAATGGATTTCACTCATGGGGTCCCTCTCTGGGTAAGCTTTTCTTTGCATTC 3237
Qy 992 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
Db 3238 TGGGTCAATTGTGACCTGTATCTCTTCTCAAGGGTCTCATGGGCGAGGAGAACCCGAACA 3297
```

```
Qy 1012 ProThrValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1031
Db 3298 CCCACCATCGTGTATTGGTCAATCTCTGGCATCTGTTTCTCTCTCTCTCTCTCTCT 3357
Qy 1032 ArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn 1051
Db 3358 AGAATTGATCTCTTCTTGATGAGTTAAAGGCCAGATACTAAACAATGTGCATCAAC 3417
Qy 1052 Cys 1052
Db 3418 TGC 3420

RESULT 11
BD236020 3851 bp DNA linear PAT 17-JUL-2003
LOCUS Materials and method for modification of plant cell wall
DEFINITION Polyaccharides.
ACCESSION BD236020
VERSION JP 2002527056-A/42.
KEYWORDS Pinus radiata (Monterey pine)
SOURCE Pinus radiata
ORGANISM Pinus radiata
REFERENCE 1 (bases 1 to 3851)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
JOURNAL BLOKBERG, L.N.
COMMENT Materials and method for modification of plant cell wall
Patent: JP 2002527056-A 42 27-AUG-2002;
GENESIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE
FORESTS LTD
OS Pinus radiata (radiata pine)
PN JP 2002527056-A/42
PD 27-AUG-2002
PF 08-OCT-1999 JP 2000575985
PR 13-OCT-1998 US 09/170862, 11-AUG-1999 US 60/148426 PI
LEONARD NATHAN BLOKBERG
PC C12N15/09, A01H5/00, C12N5/10, C12N9/00, C12N15/00, C12N5/
PC 00
CC Materials and method for modification of plant cell wall CC
FH Key Location/Qualifiers
FT source 1. 3851
FT /organism='Pinus radiata (radiata pine)'.
FEATURES
source
1. 3851
/organism='Pinus radiata'
/mol_type='genomic DNA'
/db_xref='taxon:3347'
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3851
Score: 4477.00 Matches: 839
Percent Similarity: 84.56% Conservative: 92
Best Local Similarity: 76.20% Mismatches: 104
Query Match: 79.27% Indels: 66
DB: Gaps: 16
US-10-627-132-30 (1-1052) x BD236020 (1-3851)
Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 153 ATGGAACCGACCGCGCTTGGTTCGGTTCCTCATACAGAAACGAGTTCGTGGTCATC 212
Qy 21 ArgGlyHisGluAspProLysProLeuArgAlaLeuSerGlyClnValCysGluIleCys 40
Db 213 CATGGACATGAGCGGAGCGCTTTGAAACACGCTTGAGTGGCCACGCTCTGCCAGATTGT 272
Qy 41 GlyAspGluValGlyLeuThrValAspGlyAspLeuPheValAlaCysAsnGluCysGly 60
Db 273 GCGGAGGAGCTGGGCTTTAACACAGCGCGAGCTGTTGTTGCTCTGTATGATGATGCGGG 332
```

```
QY 61 PheProValCysArgProCysTyrGluTyrGluArgGluGluGluThrGlnAsnCysPro 80
DB 333 TTTCTGTCTGTGCGCGCTATGATGACGAGACGAGAGGAATCAGTGTGCGCG 392
QY 81 GlnCysLysThrArgTyrLysArgLeuLysGlySerProArgValAlaGlyAspAsp 100
DB 393 CAGTCCAAATACCTCGTTTACAGCGTCAAAAGGAGTCCACGGGTGGAAGGTGACCAT 452
QY 101 GluGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArg 120
DB 453 GAAGAAGAGCGTTGATGACATAGAACATGAATTAATGTGGAGACTCAGCAAAAGACAG 512
QY 121 GlnLeuGluGlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMet 140
DB 513 CAG-----CAGATCACCGAGCGCATGCTCCACGAGCGCATG 548
QY 141 SerTyrGlyArgGlyProAspAspGlyAspGlyAsn-----AsnThrProGlnIlePro 158
DB 549 AGCTATGCGCGAGGTCCCGACGACGAAATTCGAGATGCTCATATCCAGAGGCTTCT 608
QY 159 -----ProIleIleThrGlySerArgSerValProValSerGlyGluPheProIle 175
DB 609 CGCAGATTCCTGTACTTGCAACCGCCACTCGGTT---GTGAGTGGGAGATTCCA--- 662
QY 176 ThrAsnGlyTyrGlyHisGlyGluVal-----SerSerSerLeuHisLysArgIleHis 193
DB 663 ACGTCATATACGACAGACAAACCAATGTGCTTGCAACCTCCGAATCTGAAGCGTGTGCAT 722
QY 194 ProTyrProValSerGluProGlySerAlaLys-----TrpAspGluLysGlu--- 210
DB 723 -----CCAAGCTCGAGCGGGAGTGAAGGATCATCATGATGCCAACAGGATAT 776
QY 211 -----ValSerTrpLysGluArgMetAspAspTrpLysSerLys 223
DB 777 GGTCTTATGCTTGGGNAAGTGTCTTGGAAGAGCGAGCGGATGTTATATAATCGAAG 836
QY 224 GlnGlyIleLeuGly-----GlyGly--- 230
DB 837 GAAAACAAATCAGGCGAGTGGATATGACGGAAGGAGATATCAATATAATGGGGGTTT 896
QY 231 -----AlaAspProGluAsp---MetAspAlaAspValAlaLeuAsnAspGluAlaArg 247
DB 897 GCACCAATAGCCCTGAAGATTATATGATCCCGATATGCCAATACCGCATGAAACGAAG 956
QY 248 GlnProLeuSerArgLysValSerIleAlaSerSerLysValAsnProTyrArgMetVal 267
DB 957 CAGCCACTGTCGCGAAAGTGCCTCCTTCAAGCAAAATAAATCCATCCGATGGTC 1016
QY 268 IleValValArgLeuValValLeuAlaPheLeuArgTyrArgIleLeuHisProVal 287
DB 1017 ATTGTAATTCGACTGATAGTGTGGGTATTTTCTCCGCTATCTGCTCCTGAATCCAGTG 1076
QY 288 ProAspAlaIleGlyLeuTrpLeuValSerIleIleCysGluIleTrpPheAlaIleSer 307
DB 1077 AAGATGATATAGGCGCTCTGGGCCACTTCTATCTGTTGTGAATCTGGTTGGCTTGTCAT 1136
QY 308 TrpIleLeuAspGlnPheProLysTrpPheProIleAspArgGluThrTyrLeuAspArg 327
DB 1137 TGGATCTTGATCAGTTTCCCAAGTGTGCTATCATCGTCGTAACGTAICTTGATCGA 1196
QY 328 LeuSerLeuArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPhe 347
DB 1197 CTGTCAATTAAGGTAGCAACAGAGCGCAACCATCAATGCTTGCACCTGTGGACCTCTTT 1256
QY 348 ValSerThrValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSer 367
DB 1257 GTGAGTACTGTAGATCCACTGAAGGAGCGCTCTTGGGTACTGCCAATACAGATATATCA 1316
QY 368 IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyVal 387
DB 1317 ATCCTTTGATGATACCTCCTGTAGCAATGTGCTCTGTATGCTCTGTATGACGAGCG 1376
QY 388 SerMetLeuThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValPro 407
```

```
DB 1377 TCGATGCTTACTTTTGAATCTCTCTGAGACCTCAGAAATTTGCCAGAAAATGGGTACCA 1436
QY 408 PheCysLysLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysVal 427
DB 1437 TTTGCAAGAAATTCGACATTTGAGCTCGCGCTCCGAAATCTATTTCTCTCAGANAAT 1496
QY 428 AspTyrLeuLysAspLysValGlnProThrPheValGlnGluArgArgAlaMetLysArg 447
DB 1497 GACTATCTGAAGGACAAATTTCAACCCACCTTTGTCAAAAGAGCGCGTGCATGAAGAGA 1556
QY 448 GluTyrGluGluPheLysValArgIleAsnAlaIleValAlaLysAlaMetLysValPro 467
DB 1557 GAATATGAAGAAATTCGAAGGTGCGCATCAATCGGTGGTGTGTCAAAGGCTCTAAAAGTGCC 1616
QY 468 AlaGluGlyTrpIleMetLysAspGlyThrProTyrProGlyValAsnAsnThrArgAspHis 487
DB 1617 AAGGAGGATGACCAATGCAAGACGCTAGCGCTTGGCTGTGTATATATACCGTGACCAT 1676
QY 488 ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHisAspThrGluGlyAsnGlu 507
DB 1677 CTTGGTATGATCCAAAGTGTCTTGGGTCAACAGTGGCGCTCGATACAGAAAGGCAATGAG 1736
QY 508 LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGluHisLysLys 527
DB 1737 CTTCTCGGCTAGTATATGTTTCTCGTGAGAGAGACCTGGTTTCCAGCATCACAAAGAG 1796
QY 528 AlaGlyAlaMetAsnAlaLeuIleArgValSerAlaValLeuThrAsnAlaProPheMet 547
DB 1797 GCCCGTGCATGAATGCTTTGGTTCGGGTTTCTGCTGTGCTCACCAATGCTCCATTTATG 1856
QY 548 LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMetCys 567
DB 1857 CTGAATCTGGATTTGTGATCACTACATTAACAATAGCAAGGCAATCAGGAAAGGCGATGTC 1916
QY 568 PheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGluArgPhe 587
DB 1917 TTTATGATGGATCTCAGTTTGGGAGAAAGTCTGTATGTCCTCAATTCCTCCAGAGATT 1976
QY 588 AspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsn 607
DB 1977 GATGGTATTGATCGCAATGACCGTTACGCCAATCGAAACACCGTATTTCTTGATATCAAC 2036
QY 608 MetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArg 627
DB 2037 ATGAAAGCTCTGGATGGAATTCAGAGGCTCTATATGTGGAACTGGATGCTGATTCAGA 2096
QY 628 ArgGlnAlaLeuTyrGlyTyrAsnProLysGlyProLysArgProLysMetValThr 647
DB 2097 AGACAAAGCTCTATATGGGTATGGGCTTCCCAAGGCGCCCAAAACCGTCCCAAGATGGTGACC 2156
QY 648 CysAspCysCysProCysPheGlyArgLysLysArgLysHisAlaLys----- 663
DB 2157 TGTGATGTCTCCCTTGTGCGGTCTCTGTAAGAAGTCTCCGAGAGAAAATAGTAGCAAG 2216
QY 664 -----AspGlyLeuProGluGlyThrAla 671
DB 2217 AAAAGTGCAGGAATCCCGACTCCCGCTTACAATCTGGACGGGATCGAGAGGAGTAGAA 2276
QY 672 AspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArg 691
DB 2277 -----GGTTATGATGCAAAAGAGCATTTGTGTATGAGCCCACTAGACTTCGAGAGGAG 2330
QY 692 PheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProPro 711
DB 2331 TTTGGCCAGTCTTTCAGCTTTTGTTCATCCACTCTGATGGAGAAATGGTGGTGTCCGCA 2390
QY 712 SerSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGlu 731
DB 2391 ACAGCAAAATCCAGCTGAATTTGTAAGGAGGCTATTCATGTCATCAGCTGTGGATGAA 2450
QY 732 AspLysThrAspTrpGlyLeuLeuGlyTyrTrpIleTyrGlySerIleThrGluAspIle 751
```

Db 2451 GACAAACCGAATCGGGAAAAGAGCTTGGATGGATCTATGATCAGTCACAGAGACATT 2510
Qy 752 LeuThrGlyPheLeuMetHisCysArgGlyTrpArgSerValTyrCysMetProLysArg 771
Db 2511 CTGACTGATTCAGATGACACACTCGAGGCTGGGGTCCATTACTGTATGCCCAACGA 2570
Qy 772 AlalaPheLysGlySerAlaProLleAsnLeuSerAspArgLeuAsnGlnValLeuArg 791
Db 2571 GCAGCATTCAAAGGCTCGCTCCAAATCAATCATACAGACCGTTTGAACCGAGTGTGGT 2630
Qy 792 TrpAlaLeuGlySerValGluLysPhePheSerArgHisSerProLeuLeuTyrGlyTyr 811
Db 2631 TGGGCTTTGGGATCAGTAGAAATTTTCATGAGCAGACATTCCCAATCTGGTATGGCTAT 2690
Qy 812 LysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 831
Db 2691 ---GGGGAGGCTCGAATGCTTGAAGATTTGCCTATATCAACACCATTTGCTATCCA 2747
Qy 832 PheThrSerLeuProLeuLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGly 851
Db 2748 TTCACCTCTCTTCCACTCATTCGCTATTCGACACTTCAGCGCTCAGTTGCTCAGTGGC 2807
Qy 852 LysPheIleMetProSerIleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMet 871
Db 2808 AATTTGTGATCCCTCAGATCAGTACTTTTGCAGAGCTTATTTTAAATAGCTCTTTTTCAT 2867
Qy 872 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrp 891
Db 2868 TCATTTTGGCATCTGTTATCTCAAAATGAGGTGGAGTGAGCATTTGAGAAATGG 2927
Qy 892 TrpArgAsnGluGlnPheTrpValIleGlyValSerAlaHisLeuPheAlaValVal 911
Db 2928 TGGCGAAATGAACAGTCTCGGTATTGGAGGGTTTCGCACATTTTTCGACGTTATT 2987
Qy 912 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 931
Db 2988 CAAGGCTGCTCAAGGTACTGGCAGGCATTTGATACAAATTTTCAGTCACCTGCCAAGGCA 3047
Qy 932 ThrGlyAspGluAspAspGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeu 951
Db 3048 TCA-----GATGACGGTGAGTTGGGGNACTGTATGCATTCAAATGGACCACTCTC 3101
Qy 952 IleProThrThrLeuLeuIleIleAsnValIleGlyValValAlaGlyIleSerAsp 971
Db 3102 ATTCCTCTACACCTGCTGTCTCAACCTTGTGGGGTGGTGTGGTGGTGGTGGTGGTGGT 3161
Qy 972 AlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 991
Db 3162 GCAATCAACATGGATTTTCAGTCATGGGGTCCCTCTCTTGGGTAAAGCTTTTCTTTGGCATC 3221
Qy 992 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1011
Db 3222 TGGGTCAATTGGCACCTGTATCTCTTCTCAAGGCTCTCATGGGCGAGACCGGACCA 3281
Qy 1012 ProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1031
Db 3282 CCCACCATCGTGGTATTGTGTCATCTGTGGCATCTGTTTCTCTCTTTCTGGGTA 3341
Qy 1032 ArgIleAspProPheIleValArgThrLysGlyProAspValArgGlnCysGlyIleAsn 1051
Db 3342 AGAATTGATCTTTCTTGGTAAAGTTAAAGGCCCAAGATATAAACAATTTGGCATCAAC 3401
Qy 1052 Cys 1052
Db 3402 TGC 3404
AC140546 106962 bp DNA linear PLN 09-MAR-2004
LOCUS Medicago truncatula clone mth2-35119, complete sequence.
DEFINITION AC140546
ACCESSION AC140546
VERSION AC140546.10 GI:45268058
KEYWORDS HTG.

SOURCE
ORGANISM

Medicago truncatula (barrel medic)

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.REFERENCE
AUTHORS

1 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

Medicago truncatula BAC Clone mth2-35119

TITLE

Unpublished

JOURNAL

2 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.REFERENCE
AUTHORS

Direct Submission

Submitted (25-FEB-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

TITLE

3 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

JOURNAL

Direct Submission

Submitted (30-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USAREFERENCE
AUTHORS

4 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

TITLE

Direct Submission

Submitted (16-OCT-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USAREFERENCE
AUTHORS

5 (bases 1 to 106962)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.

TITLE

Direct Submission

Submitted (09-MAR-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT

On Mar 9, 2004 this sequence version replaced gi:37693616.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES

source

1. 106962

/organism="Medicago truncatula"

/mol_type="genomic DNA"

/db_xref="taxon:3880"

/clone="mth2-35119"

/clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 106962
Score: 4247.50 Matches: 863
Percent Similarity: 60.86% Conservative: 90
Best Local Similarity: 55.11% Mismatches: 83
Query Match: 75.20% Indels: 534
DB: 8 Gaps: 21

US-10-627-132-30 (1-1052) x AC140546 (1-106962)

Qy

1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20

Db

26794 ATGGAAGCCAGCCGCGGACTAGTCGCTGGTTCTCAACACCGAAGCAGCTTGTTCATT 26853

Qy

21 ArgGlyHisGluAspPro----- 26

Db

26854 CATGGCCATGAGAGAGGT-ATAATAAACTTATAACCAAGTGGTGGTAGAGATATT 26912

Qy

27 -----LysProLeu 29

Db	26913	GTTTGTTTGTATTATGTTGATGTTTGTGTTGTTTGTGTCAGCAGCCATTTG	26972
QY	30	ArgAlaLeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeuThrValAsp	49
Db	26973	AAAGAACTTGGAGCGTCAAGTGTGCGAGATATGTGGTGTATGATGTCTTACTGTGGAT	27032
QY	50	GlyAspLeuPheValAlaCysAsnGluCysGlyPheProValCysArgProCysTyrGlu	69
Db	27033	GGAGATTTTGTGTCATGTAAACGAGTGTGTTTTCGGTGTGCGGCTTGTCTATGAG	27092
QY	70	TyrGluArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyrLysArgLeu	89
Db	27093	TATGAAGAAGGAGGGGAGACAACCTTGTCTCAGTGCACAGACATACAGAGCTCTC	27152
QY	90	-Lys-	90
Db	27153	AAAGGTAGAATTATATATGTAACAAGTATTTTATAGCCTTGTGTTATCATGCAATAA	27212
QY	90	-	90
Db	27213	GTAAACATGAACCTTACATGATACAGTTTATAGATGACTGTGATATTAATTTGATCGAT	27272
QY	91	-GlySerProArgValAlaGlyAspAspAs	100
Db	27273	TAACCTAAATGCATATACTAATATGATGTTTAGGGGCCCTCGGTTGAGGAGACATGA	27332
QY	100	pGluAspIleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnAr	120
Db	27333	TCAGGAGGATGTGGACGATATTGAACATGAATTCAAAATTGAAGACAA-	27381
QY	120	gGlnLeuGluGlyAsnMetGlnAsn--SerGlnIleThrGluAlaMetLeuHisGlyAr	139
Db	27382	-ATGAACAATCATGATCATCTCTGCTGAGGCCATCTGCATCGGAA	27425
QY	139	gMetSerTyrGlyArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProPr	159
Db	27426	GATGAGCTATGAAGAGGTCTGAAGAT-----GATGAAAATGCACATCTCCAGC	27476
QY	159	oilelleThrGlySerArgSer-----	166
Db	27477	TGTATTGTGTGTGTGCTCGGNAATGTAAGAAATTTTCCACCATTTCTAGTCCATCTTA	27536
QY	167	-ValProValse	170
Db	27537	ACTCGATAGAACCTGTGAATTCATACAACTAAATTTGCTTTTATACATGCATCGAGTGA	27596
QY	170	rGlyGluPheProIle--ThrAsnGlyTyrGlyHisGlyGluValSerSerLeuHi	189
Db	27597	TGTGAGTTCCTCAATATCATCTCATAGTTATGGGAG---CAGATGCTATCTTCACTGCA	27653
QY	189	sLysArgIleHisProTyrProValSerGlu-----	199
Db	27654	TAAAGAGTGCATCCATATTCGCTTCATGTTCTCGTAGAGTAGTTATAGATTATC	27713
QY	199	-	199
Db	27714	ACTTTTCTGGTTTCATTTGCCAAAACGGATTATGCCAAAATTGGTTTAAATTTCTTGTAAAG	27773
QY	199	-	199
Db	27774	TCCAAATCTTACTAGTTATACAGCCCTTTCTTTTGGCAGTGAATAAAATACACCCCTG	27833
QY	199	-	199
Db	27834	AAATGCTTTTCTTAGCTTACTACACATGGTGGATATGATTATATATGTGGCATGCTGT	27893
QY	200	-----ProGlySerAlaLysTyrAspGluLysGlyValSerTyrLysGluArgMe	217
Db	27894	GATAAAAACAGGAAGTGCAGGCTGGATGAAAGGAGAGAAATGATCATATATAGAAAT	27953
QY	217	thAspThrLysSerLysGlnGlyIleLeuGlyGlyAlaAspProGluAspMetAs	237
Db	27954	GGATCACTGGAAATTCAGCAAGGAAATTTGGGACCTGAACCTGAT---	28010
QY	237	palAspVal-----	240
Db	28011	TGCAAAACATGATGCTTTTTCAGTACTAAATGCATTTCTCTGACAAATTTAAATCTTAATGAAA	28070
QY	240	-	240
Db	28071	TGATAGTAAATTTTGAAGATCATAAATTAATTAATTTAGTTACTACTATACAAATGTG	28130
QY	241	-----AlaLeuAsnAspGluAlaArgGlnProLeuSerArgLysValSerIleAlase	258
Db	28131	AACCCATGCAGG-TCAGATGAAGCAAGCAACCACTGTCAAGAGGTACCAATAGATCATC	28189
QY	258	rSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPhePh	278
Db	28190	TAGCAAAATCAATCCATAGATGCTGATCGTGACCGCTTGTATTCTTTGGCTTCTT	28249
QY	278	eLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTyrLeuValSerIle	298
Db	28250	CCTCCGATACAGACTTATGAATCCAGTACATGATCAATGGGGCTATGGCTAACTTCAAT	28309
QY	298	eileCysGluIleTyrPheAlaIleSerTyrIleLeuAspGlnPheProLysTyrPhePr	318
Db	28310	TATATGTGAATCTGGTTTGTATATCATGATCTTGATCTGCCCAATGGTATCC	28369
QY	318	oileAspArgGluThrTyrLeuAspArgLeuSerLeu-----	330
Db	28370	TATTGATAGAGAAACATACCTTGCCTGCTTCTCACTCAGGTAAGGGACCTTCCCTCCCTG	28429
QY	331	-	331
Db	28430	ATTGTTGAAATACCACTGTTTAAATTAATTTATATCTCTGAACAAAGAAAGTTGTTTATCAG	28489
QY	331	gTyrGluArgGluGlyGluProSerLeuSerAlaValAspLeuPheValSerThrVa	351
Db	28490	GATGAGCGTGAAGCGCAACCAATATGCTTGTCTCTGTAGATGTGTTGTAGTACTGT	28549
QY	351	laspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAlaVa	371
Db	28550	GGATCCCTTGAAGGAACCTCTCTGATACAGCAACACAGTTCTTCAATCTTGGCAAT	28609
QY	371	laspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuTh	391
Db	28610	GGACTACCCCATGTATAAGATATCATGTACATTTCTGATCAGCGAGCTTCAATGTGATC	28669
QY	391	rPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTyrValProPheCysLysVal	411
Db	28670	ATTGGAACCTTTCGGAACCGGAGATTTGCTAGGAAGTGGGTACCATTTTGTAAAGAA	28729
QY	411	sPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLeuLys	431
Db	28730	ATTTTGTATAGAACCTCTGTCACACAGATGTACTTCTCTGAGAAAATTTGACTACTAAA	28789
QY	431	sAspLysValGlnProThrPheValGlnArgAlaMetLys-----	446
Db	28790	GGCAAGGTGAGCCAACTTTTGTCAAGAAACGTGGTCCATGAAAGGTTTGTATTAAATCT	28849
QY	446	-	446
Db	28850	ATTAGATTTGTAAATAAATAAGTCTGAAATCAGTTTAAAGCTACTCACCAGCTA	28909
QY	447	-----ArgGluTyrGluGluPheLysValArgIleAsnAlaLeuVa	460
Db	28910	TTAATTGGATGCATGGCAGAGAGATATGAAGATTTTAAGGTTAGATCAATGCATCTGT	28969
QY	460	lAlaLysAlaMetLysValProAlaGluGlyTyrIleMetLysAspGlyThrProTyrPr	480
Db	28970	GGCAAAAGCTCAGAAGGTTTCTGCGAGGAGGTGGATATGAGGATGGGACACCATGGCC	29029
QY	480	oGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGl	500
Db	29030	AGGAAACAATACTAAGGATCATCTCTGGTATGATTCACGCTCTTCTTGGTCACAGTGGAGG	29089

```
Qy 500 yHieAspThrGluGlyAenGluLeuProArgLeuValTyrValSerArgGluLysArgPr 520
Db 29090 TCATGATAGTAGAGAAACACGCTCTCTCGCTGTTTATGATCTAGAGAGAAAGGCC 29149
Qy 520 oGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeu----- 534
Db 29150 AGGATTTCAACACACCAAGAAAGCGGTGCCATGAATGCTCTGCTAGAGTTACTTTGAACA 29209
Qy 534 ----- 534
Db 29210 AACAGATTTGTTATTTTATGATGTCATTCACCTTTCTGAAACACACTTTACTTGACTC 29269
Qy 535 ----- 546
Db 29270 ATCTAAATAATTATGTAATCTCCAGGTACGGGTCTCTCGACTGCTTACAAATGCTCCTTT 29329
Qy 546 eMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaIleArgGluAlaMe 566
Db 29330 TATGCTGAATCTGGATTGTGATCATATATCAATAACACAGCAAGGCTGTCCGAGAGGCCAT 29389
Qy 566 tCysPheLeuMetAspProGlnValGlyArgLysValCysTyrValGlnPheProGlnAr 586
Db 29390 GTGCTTCTTGATGGACCCCAACAGGGAAGAAAGTCTGCTATGTCCAGTTTCTCTCAAG 29449
Qy 586 gPheAspGlyIleAspValHisAspArgTyrAlaAsnArgAsnThrValPhePheAspIl 606
Db 29450 ATTGATGGCATTGATGGCATGATCATGATATGCTAATAGGAACACAGTTTCTTTGATGT 29509
Qy 606 e----- 606
Db 29510 -AAGTCACTGCAAGAAAGTCAACATCAACAGCATGATCTCTTTTGAAGCAATTGATGT 29568
Qy 606 ----- 606
Db 29569 TTTTGTAGTAGCAAGCTAATATATATATACCAATTCCTCTGTTTCTGTCATGGCAGAT 29628
Qy 607 -AsnMetLysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPh 626
Db 29629 TAACATGAAGGGTCTAGATGGTATTCAGGGTCTCTGTATGTGGCAGAGGGTGTGTTT 29688
Qy 626 eArgArgGlnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetVa 646
Db 29689 TAGAAGGCAGGGCTTTATATGGATATAATCTCTTAAGGGTCCCAAGCGTCCAAAATGGT 29748
Qy 646 lThrCysAspCysCysProCysPheGlyArgLysLysArg--LysHisAlaLysAsp- 664
Db 29749 AAGTTGTGATGTCGCCATGTTTGGAAAGGCGCAAGAGGTTAAGCAATGCAATGAACGA 29808
Qy 665 -----GlyLeuProGlu----- 668
Db 29809 CCCAATGGAGAGCTGCAGGCTAAG-AGGTTCTAATTTTACAAATAGTCAGCAATCTT 29867
Qy 669 -----GlyThrAl 671
Db 29868 GATACTAAATCCTTATACAAATGGATTTTATGTTTCACTGTTCTTAATCTCTGTACATGT 29927
Qy 671 aspMetGlyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysAr 691
Db 29928 CCATATAGGAATGGGAAGATGACAAAGAGTTACTGATGTCCTCCAGATGAATTTTGAAGAA 29987
Qy 691 gPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyGlyValProPr 711
Db 29988 ATTTGGGCAGTCATCGATTTTGTGACTTCTAGTCTTGTATGGGAAGGGTGTGATCTCC 30047
Qy 711 oSerSerSerProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGl 731
Db 30048 TTTCGTCAAGTCCAGCAAGCCAACTTAAAGAAAGCCATTATGTAATCAGTTGCGGATATGA 30107
Qy 731 uAspLysThrAspTyrGlyLeuGlu----- 739
Db 30108 AGATAAAACTGATGGGGGATTTGAGGTAAACCAACATTTGGCAGCGCTAGGCTCTCAATTG 30167
Qy 739 ----- 739
Db 30168 ATATTTTACAATAGTTGTTTCCAAATAAGAAATCAGCTCATCCATATTTTGTTCCAA 30227
Qy 740 -----LeuGlyTyrIleTyrGlySerIleThrGluAspIleLeuThrGlyPheLysMe 757
Db 30228 ATTGCAGCTTGGTGGATTTATGGATCTATTACAGAGGATATTTTACAGGCTTTAAGAT 30287
Qy 757 tHisCysAspGlyTyrArgSerValTyrCysMetProLysMetProLysArgAlaAlaPheLysGlySe 777
Db 30288 GCATTGCCGTGGTGGAGATCCATTTACTGTATGCCAAAGAGAGTAGCATTTCAAGGTAC 30347
Qy 777 rAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrPheAlaLeuGlySerVa 797
Db 30348 TGCTCCTATCAACTTGTTCAGATAGACTCAACACAGGTGCTTCGTTGGGCTCTTGCTCCAT 30407
Qy 797 lGluIlePhePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817
Db 30408 TGAGATCTCTTTCAGTCATCAATTGCCCTTTATGGTATGGCCCAAGGAAGGAAAGCTGAA 30467
Qy 817 sTrpLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLe 837
Db 30468 GTGGCTGGAGCGATTGGCTATGCMAACACACACCGGTACCCCTTCACTCCATACCTCT 30527
Qy 837 uLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro-- 856
Db 30528 AGTTGCCCTACTGTATTCTTCAGCTGTCTGCTTACTCACTGACAAATTCATCATGCCACC 30587
Qy 856 ----- 856
Db 30588 GGTAACATTTTATTTCTTTTATCTTCTGTTTAGAAAATGAGTTAAGACACGACGAGCTTT 30647
Qy 856 ----- 856
Db 30648 TGAATTTATTAATTTTTTGACAAACATTACAATATTTTCAATATATTTACCTTCAAAAGCACA 30707
Qy 856 ----- 856
Db 30708 AGAATGATCAGCATCTTTTATGTATGTAGTAGGAATTCAACTTTTAAATCCATAACAGC 30767
Qy 856 ----- 856
Db 30768 TAGAAATGCTAGGCATCAATTGTTCAAGGATGCTCAGAGAAATTCATGCAAGCCTTTTTCAC 30827
Qy 857 -----SerIleSerThrPheAlaSerLeuPhePheIl 867
Db 30828 TCAATTTTTTTTTTAATTTGTTTCTTCTTCAGATAAGCACTTTGCTAGTTTGTACTTGT 30887
Qy 867 eAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValSe 887
Db 30888 TGCTCTCTCTCTTCAATCATGCAAGCGGCACTTCTTGAGTTGAAATGGAGTGGAGTCAG 30947
Qy 887 rIleGluGluTyrTrpArgAsnGluGlnPheTyrValIleGlyGlyValSerAlaHisLe 907
Db 30948 CATTGAGGAATGGTGGAGAAATGAGCAATCTGCGTCAITGGTGTATCAGCACATCT 31007
Qy 907 uPheAlaValValGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrVa 927
Db 31008 CTTTGTGTATACAGGCTTCTGAAAGTCTTGTCTGGAATTCATGACACAACTTCACTGT 31067
Qy 927 lThrSerLysAlaThrGlyAspGluAspGluPheAlaGluLeuTyrAlaPheLysTr 947
Db 31068 TACATCCAAAGCAACA-----GATGATGAGGAATTTGGAGAATTTGACGCCATCAAGTG 31121
Qy 947 pThrThrLeuLeuLeuProProThrThrLeuLeuLeuIleLeuValIleGlyValValAl 967
Db 31122 GACCACCTCTCTAATCTTCCACAGCATCTCTTAATAATCAATATGTTGGGGTGTGTC 31181
Qy 967 aGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLe 987
Db 31182 TGGATCTCAGACCCCATTAACATGTTTACCAATCATGGGACCTCTTATTTGGAAAAC 31241
Qy 987 uPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyAr 1007
```



```
Db 104057 TAAAGAGTGATCCATATTCGGCTTCGTGATTCCTCGTAAGATGTAGTTTATAGATTTATC 104116
|||:::||||| |||:::
Qy 199 -----
Db 104117 ACTTCTTGGTTCAATTCGAAACCGGATATGCCAAAAATGGTTAAATTTCTTGCTAAAG 104176
|||:::||||| |||:::
Qy 199 -----
Db 104177 TCCAAATCTTATAGTTATACAGGCCCTTCTTTTGGCACTGATTAATAATACAAACCCGTG 104236
|||:::||||| |||:::
Qy 199 -----
Db 104237 AAATGCTTTTCTAGCTTACTACAACATGGTGGATATGATTAATAATATGTGGCATGCTGT 104296
|||:::||||| |||:::
Qy 200 -----ProGlySerAlaLysTrpAspGluLysGluValSerTrpLysGluArgMe 217
|||:::||||| |||:::
Db 104297 GATRAAAACAGCAAGTGCAGGCTGGGATGAAAGGAGAGAGATGGATCATATGATAGAA 104356
|||:::||||| |||:::
Qy 217 tAspAspTrpLysSerLysGlnGlyIleLeuGlyGlyAlaAspProGluAspMetAs 237
|||:::||||| |||:::
Db 104357 GGATGACTGGAAATTCGACGAGGAATTTGGACCTGAACCTGAT---GAAGATCTAGA 104413
|||:::||||| |||:::
Qy 237 pAlaAspVal----- 240
|||:::||||| |||:::
Db 104414 TGCAACATGATGCTTTTCAGTACTAAATGCAATTCCTCGACAATTTTAATTTCTAATGAAA 104473
|||:::||||| |||:::
Qy 240 -----
Db 104474 TGATAGTAATATTTGAAAGATCATAAATTAATTAATTTAGTTACTACTATACAATGTG 104533
|||:::||||| |||:::
Qy 241 -----AlaLeuAenAspGluAlaArgGluProLeuSerArgLysValSerIleAlaSe 258
|||:::||||| |||:::
Db 104534 AACGATGAGG-TCAGATGAAGCAAGACACCACTGTCAAGGAGGTACCAAATAGCAATC 104592
|||:::||||| |||:::
Qy 258 rSerLysValAenProTrpArgMetValIleValArgLeuValValLeuAlaPhePh 278
|||:::||||| |||:::
Db 104593 TAGCAAAATCAATCCATATAGATGGTGTCTGTGGCAGCGCTGTATTTCTTGGCTCTT 104652
|||:::||||| |||:::
Qy 278 eLeuArgTrpArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSerI 298
|||:::||||| |||:::
Db 104653 CCTCCGATACAGACTATGAAATCCAGTACATGATGCAATGGGGCTATGGCTAACTTCAAT 104712
|||:::||||| |||:::
Qy 298 eIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPhePr 318
|||:::||||| |||:::
Db 104713 TATATGTGAATCTGGTTTGTATATCATGATGATCTTTCATCGTTTCCCAATGGTATCC 104772
|||:::||||| |||:::
Qy 318 oIleAspArgGluThrTrpLeuAspArgLeuSerLeu----- 330
|||:::||||| |||:::
Db 104773 TATGATAGAGAAACATACCTTGACCGCTTTTCACTCAGGTAAGGGACCTTCCCTCCTG 104832
|||:::||||| |||:::
Qy 331 -----Ar 331
|||:::||||| |||:::
Db 104833 ATTTGTAATAATCAGTGTTTAAATATTTATATCTCTGAACAAAGAAAGTTGTTTATCAG 104892
|||:::||||| |||:::
Qy 331 gTrpGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerThrVa 351
|||:::||||| |||:::
Db 104893 GTATGAGCTGAAGCGAAGCAACCAATATGCTTCTCTGTAGATGTGTTGTAGTACHTGT 104952
|||:::||||| |||:::
Qy 351 lAspProLeuLysGluProProLeuValThrAlaAenThrValLeuSerIleLeuAlaVa 371
|||:::||||| |||:::
Db 104953 GGATCCCTTGAAGGAACCTCTCTGAATACAGCCCAACACAGTCTTTTCAATCTTGGCAAT 105012
|||:::||||| |||:::
Qy 371 lAspTrpProValAspLysValSerCysTrpValSerAspAspGlyAlaSerMetLeuTh 391
|||:::||||| |||:::
Db 105013 GGACTACCCCAATGATATGATATCATGCTACATTTTCTGATGACGAGCTTCAATGTGAT 105072
|||:::||||| |||:::
Qy 391 rPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLysAl 411
|||:::||||| |||:::
Db 105073 ATTTGAAGCCCTGTGGAACCGCAGAGTTTCTAGGAAGTGGTACCATTGTTGTAAGAA 105132
|||:::||||| |||:::
Qy 411 sPheGlyIleGluProArgAlaProGluPheTrpPheSerLeuLysValAspTrpLeuLy 431
|||:::||||| |||:::
Db 105133 ATTTTTCATAGAAACCTCGTGCAACGAGATGTACTTCTCTGAGAAAAATTGACTATCTAAA 105192
|||:::||||| |||:::
Qy 431 sAspLysValGlnProThrPheValGlnGluArgAlaMetLys----- 446
|||:::||||| |||:::
Db 105193 GGCAAAAGGTGAGCAACACTTTTGTCAAAGAACGTGGTGCCATGAAGGTTTGTATTAATCT 105252
|||:::||||| |||:::
Qy 446 -----
Db 105253 ATTAGAATTGTTGAAAAAATGAAGTCTCGAAATCAGTTTAAAGCTTACTCACCGACTA 105312
|||:::||||| |||:::
Qy 447 -----ArgGluTrpGluGluPheLysValArgIleAenAlaLeuVa 460
|||:::||||| |||:::
Db 105313 TTAATTGGATTCATGGCAGAGAAATATGAAGAGTTTAAAGGTTAGGATCAATGCACCTGT 105372
|||:::||||| |||:::
Qy 460 lAlaLysAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpPr 480
|||:::||||| |||:::
Db 105373 GGCAAAAGCTCAGAAGGTTCTCGAGAGGGTGATTTATCGAGATGGGACCATGGCC 105432
|||:::||||| |||:::
Qy 480 oGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGl 500
|||:::||||| |||:::
Db 105433 AGGAACAATACTAAGGATCATCTCGTATGATTCAGTCTTTCTTGGTCACAGTGAGG 105492
|||:::||||| |||:::
Qy 500 yHisAspThrGluGlyAsnGluLeuProArgLeuValTrpValSerArgGluLysArgPr 520
|||:::||||| |||:::
Db 105493 TCATGATAGTGAAGAAACAGCTTCTCGCTTGTGTTATGTATCTAGAGAGAAAGGCC 105552
|||:::||||| |||:::
Qy 520 oGlyPheGlnHisLysLysValAlaGlyAlaMetAsnAlaLeu----- 534
|||:::||||| |||:::
Db 105553 AGGATTTCAACACCAAGAAAGCCGTGCATGATGCTCTGGTAGAGTTACTTGAACA 105612
|||:::||||| |||:::
Qy 534 -----
Db 105613 AACAGATTTTGTATTTTATGATGTCATTCACTTCTTGAACCCACACTTACTTGACTC 105672
|||:::||||| |||:::
Qy 535 -----IleArgValSerAlaValLeuThrAsnAlaProPh 546
|||:::||||| |||:::
Db 105673 ATCTAATAATTTATGTAACTTCCAGGTACGGGTCTCTGCGATGTGTTACAAATGCTCCTT 105732
|||:::||||| |||:::
Qy 546 eMetLeuAsnLeuAspCysAspHisTrpIleAsnAsnSerLysAlaIleArgGluAlaMe 566
|||:::||||| |||:::
Db 105733 TATCTGAACCTGGATGTGATCAATATATCAATPACAGCAAGGCTGTCCGAGAGGCCAT 105792
|||:::||||| |||:::
Qy 566 tCysPheLeuMetAspProGlnValGlyArgLysValCysTrpValGlnPheProGlnAr 586
|||:::||||| |||:::
Db 105793 GTGCTTCTGTATGACCCCAACAGGAAAGAAAGTCTGCTATCTCCAGTTTCTCTCAAAG 105852
|||:::||||| |||:::
Qy 586 gPheAspGlyIleAspValHisAspArgTrpAlaAsnArgAsnThrValPhePheAspIl 606
|||:::||||| |||:::
Db 105853 ATTCGATGGCATTTGATGCGATGATCGATATGCTTAATAGGAACACAGTTTTTCTTTGATGT 105912
|||:::||||| |||:::
Qy 606 e----- 606
|||:::||||| |||:::
Db 105913 -AAGTCACTGCAGAAAGTCAACATCAACAGCATGATCTTCTTTGAGCAATTTGATGT 105971
|||:::||||| |||:::
Qy 606 -----
Db 105972 TTTTGTAGTAGTCAAGCTAATATATATTACCAATTCCTCTTGTTCGTCTATGCGCAGAT 106031
|||:::||||| |||:::
Qy 607 -AsnMetLysGlyLeuAspGlyIleGlnGlyProValTrpValGlyThrGlyCysValPh 626
|||:::||||| |||:::
Db 106032 TAACATGAAGGCTCTAGATGGTATTCAGGGTCTCTGTGTATCTCGGCACAGGGTGTGTTTT 106091
|||:::||||| |||:::
Qy 626 eArgArgGlnAlaLeuTrpGlyTrpAsnProProLysGlyProLysArgProLysMetVa 646
|||:::||||| |||:::
Db 106092 TAGAAGGCGAGCTTTATATGATATATCTTCTTAAGGTCCTCAAGCGCTCCAAAAATGCT 106151
|||:::||||| |||:::
Qy 646 lThrCysAspCysCysProCysPheGlyArgLysArg---LysHisAlaLysAsp-- 664
|||:::||||| |||:::
Db 106151 AAGTTGTGATTTGCTGCCCATGTTTGGAAAGCGCAAGAGGTTAAGCATGCAATGAACGA 106211
|||:::||||| |||:::
Qy 665 -----GlyLeuProGlu----- 668
|||:::||||| |||:::
Db 106212 CGCAATGGAGAGCTGCGAGGCCCTAAG-AGGTTCTAAATTTTACAATAGTCAGCAATCTT 106270
|||:::||||| |||:::
```

```
QY 669 -----GlyThrAl 671
Db 106271 GATACTAAATCCCTTATPACAAATTTGATTTTCATGTTCCACCTGTTCTAATCTCTGTGCATCATGT 106330
QY 671 aAspMetGlyValaAspSerAspLysGluMetSerHisMetAsnPheGluLysVal 691
Db 106331 CCATATAGAGANTGGAGATGACAAAGATTTACTGTGTCCAGATGAATTTTGGAGAGAA 106390
QY 691 gPheGlyGlnSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProPr 711
Db 106391 ATTGGGCAGTCATCGATTTTGTGACTTCAGTCTTTGATGGAGAGGGTGTGACTCTCC 106450
QY 711 oSerSerSerProAlaAlaLeuLeuLysGluAlaLleHisValLleSerCysGlyTyrGl 731
Db 106451 TTTCGCAAGTCCAGCAGCAAGCAACTTAAGAAGCCATTCATGAATCAGTTGGCGATATGA 106510
QY 731 uAspLysThrAspTTPGlyLeuGlu----- 739
Db 106511 AGATAAACTGAATGGGGGATTTGAGGTAAACCAATGGGAGCCCTAGGCTCTCAATGG 106570
QY 739 ----- 739
Db 106571 ATATTTTACAATAGTTGTTTACAAATAAGAAATCAGCTCATCCATTTATTTTGTTCCAA 106630
QY 740 -----LeuGlyTTPilLysGlySerIleThrGluAspLleLeuThrGlyPheLysMe 757
Db 106631 ATTGCAGCTTGGTGGATTTATGGATCTATTACAGAGGATATTTTAAACAGGCTTTAAGAT 106690
QY 757 thsCysArgGlyTTPArgSerValTyrCysMetProLysArgAlaAlaPheLysGlySe 777
Db 106691 GCATTGCCGTGGTGGAGATTCATTTACTGTATGTCAAAGAGAGTAGCATTCACAGGGTAC 106750
QY 777 rAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTTPAlaLeuGlySerVa 797
Db 106751 TGCTCTCATCACTTGTGCAGATAGACTCAACCCAGGTGCTTCGTGGGCTCTTGGCTCCAT 106810
QY 797 lglullePheSerArgHisSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLys 817
Db 106811 TGAGATCTCTCTTCAGTCATCATTTGCCCTTTATGGTATGCCACAGGAAGGAAGTCAAA 106870
QY 817 sTPLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProPheThrSerLeuProLe 837
Db 106871 GTGGCTGGAGCGGATTTGCCCTATGCAAAACACACCGGTGACCCCTTCACCTCCATCCCTCT 106930
QY 837 uLeuAlaTyrCysThrLeuProAlaValCysLeuLeuThrGlyLysPheIleMetPro-- 856
Db 106931 AGTTGCTACTGTATTTCTCCAGCTGCTGCTTACTCACTGACAAATTCATCATGCGCAC 106990
QY 856 ----- 856
Db 106991 GGTAACATTTTATTCTTTTATCTTCTGTTAGAAAAATGAGTTAAGACACGACGAGCTTT 107050
QY 856 ----- 856
Db 107051 TGAANAATATAATTTTGTGACAAACATTAACAATTTTATATATTTACCTTCAAAAGCACA 107110
QY 856 ----- 856
Db 107111 AGAATCATCAGCATCTTTTATGTATGTATAGTAGGAATTCACCTTTTAATTCATACAGC 107170
QY 856 ----- 856
Db 107171 TAGAAAATGTAGGCATCATTTGTTCAAGGATGCTCAGAGAATTCATGCCAAGCCTTTTATC 107230
QY 857 -----SerIleSerThrPheAlaSerLeuPhePheIle 867
Db 107231 TCAATTTTATTTTAAATTTGTTTCATTTCTTCAGATAAGCACTTTTGTAGTTGTACTTTGT 107290
QY 867 eAlaLeuPheMetSerIlePheAlaThrGlyIleLeuGluMetArgTTPSerGlyValSe 887
Db 107291 TGCTCTCTCTCTTCAATCATCGCAACGGGCATTTCTTGAGTTGAAATGAGTGGAGTCAAG 107350
```

887 rIleGluGluTTPTrpArgAsnGluGlnPheTrpValIleGlyGlyValSerAlaHisLe 907
107351 CATTGAGGAATGGTGAGAAATGAGCAATTTCTGGGTCAATTTGGTGTATACGACCATCT 107410
907 pPheAlaValValGlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrVa 927
107411 CTTTCTGTGCATPACAAAGGCTTCTGAAAGTTCTTCTGAAATGTGACAAACTTCACTGT 107470
927 tThrSerLysAlaThrGlyAspGluAspGluPheAlaGluLeuTyrAlaPheLysTr 947
107471 TACATCCAAAGGCAACA-----GATGATGAGGAAATTTGAGAAATTTGACCCATCAAGTG 107524
947 pThrThrLeuLeuIleProProThrThrLeuLeuLeuIleLeuValIleGlyValValAl 967
107525 GACCACTCTCTAAATTTCTCCAAACGACTATCTATAATCAATATTTGTTGGGTTGTTGC 107584
967 aGlyIleSerAspAlaIleAsnAsnGlyTyrGlnSerTrpGlyProLeuPheGlyLysLe 987
107585 TGGAAATCTCAGAGCGCATTAACATGGTTACCAATCATGGGACCTCTATTTGGAAAACT 107644
987 uPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyAr 1007
107645 CTTCTTTCTCTCTGGTGATTTGCCATCTATATCAATCTCTTAAAGGTTTGTATGGTCTG 107704
1007 gGlnAsnArgThrProThrValValValIleTrpSerIleLeuLeuAlaSerIlePheSe 1027
107705 GCAAAACCGCACCCCCACCATTTGTTGATATGGTCAGTGTGTTGGCTTCCATTTTCTC 107764
1027 rIleuLeuTrpValArgIleAspProPheIleValArgThrLysGlyProAspValArgGl 1047
107765 TTTGTTATGGGTGAAGATTTGACCCATTCGTATGAAAACTAAGGACCTGATACCAAGCT 107824
1047 nCysGlyIleAsnCys 1052
107825 ATGGGAATCAACTGT 107840

RESULT 14
AX048245
LOCUS AX048245
DEFINITION Sequence 1 from Patent WO0070058.
ACCESSION AX048245
VERSION AX048245.1 GI:11877016
KEYWORDS
SOURCE Arabidopsis sp.
ORGANISM Arabidopsis sp.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
REFERENCE
AUTHORS Turner,S. and Taylor,N.
TITLE Plant cellulose synthase genes
JOURNAL Patent: WO 0070058-A 1 23-NOV-2000;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
source
1. .7234
/organism="Arabidopsis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:29726"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 7234
Score: 4224.50 Matches: 856
Percent Similarity: 61.35% Conservative: 87
Best Local Similarity: 55.69% Mismatches: 90
Query Match: 74.80% Indels: 508
DB: 6 Gaps: 20
US-10-627-132-30 (1-1052) x AX048245 (1-7234)

QY 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
Db 92 ATGGAAGCTAGCGCGGCTTTGTGCGCGGTTCTCATAAACCGTAATGAAGTAGTCGTCATT 151

Qy 21 ArgGlyHisGluAsp----- 25
Db 152 CACAACCATGAAGAGGTTTTTTCACATTTACTTTTCTCATCACTTACCAGAAAGTTTTTTTT 211
Qy 25 ----- 25
Db 212 TACCAACGCTAGTAAATAATTAATGTCATTTTTTCGTTTTATTTGGTTACTATTCTTAAA 271
Qy 25 ----- 25
Db 272 AATTCGTATGTTTGGGAATAAAATATATGATCAATTTTTTTTAAATCATCTTCTTATTATGA 331
Qy 25 ----- 25
Db 332 GACAAAAATTAATCTGTAATCTGTAGTTGCAATAATGTTGAGAAAAATTCATATCTT 391
Qy 25 ----- 25
Db 392 TGTAGCAACATATAATATTTTGTGTTGTAATTAATTAAGTTGAGAAGTCAGGTTTAAACCAT 451
Qy 25 ----- 25
Db 452 TTAATCGCTGTCAATTTTTTTTATTATCTTTTACTTCAATGAGATCTTTTTTTTGTGATTAG 511
Qy 26 -----Pro 26
Db 512 AAATTTCCAGGTTTTTATTTTCGCTTTTAAAGAACTTAAACTTTTTTGTGTAATTAACAGCCA 571
Qy 27 LysProLeuArgAlaLeuSerGlyGlnValCysGluIleCysGlyAspGluValGlyLeu 46
Db 572 AAGCCTCTGAAGAATCTAGATGGACAAATCTGTGAGATATGTGAGATCAGATGGTTTA 631
Qy 47 ThrValAspGlyAspLeuPheValAlaCysAsnGluCysGlyPheProValCysArgPro 66
Db 632 ACAGTAGAAGAGGACCTCTTCGTAGCTTGCAATGAGTGTGTTTTTCCGGCGTAGACCT 691
Qy 67 CysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyr 86
Db 692 TGCTATGATGACGAGAAGAAGAAACACAAACTGCTCCTCAGTGTAAAGACTCGTTAC 751
Qy 87 LysArgLeuLys----- 90
Db 752 AAGCGTCTCAGAGTAAGTTATTATTAACTCTCCCTCTGCTTGTGTTGTTGCGACGAA 811
Qy 90 ----- 90
Db 812 TGCCTCTATGAAATTTAAAGGCTGTTCTCTTTTGTAGTTGAACCTTGAGAGTAATG 871
Qy 91 -----GlySerProArgValAlaGlyAspAspGluAsp 103
Db 872 ATCTGTTTTTTGGTTTCTGAAGGAAGCCCAAGAGTGGAGGAGATGAAGACGAAGAAGAT 931
Qy 104 IleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArgGlnLeuGlu 123
Db 932 ATTGATGATATGATGATGAATTAATTCGAAACATGAACAAGAT----- 976
Qy 124 GlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGly 143
Db 977 -----AAGCATAAAGCATCTCTGCTGAGGCTATGCTTTATGCGAAATAGCTATGGA 1027
Qy 144 ArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIleIleThrGly 163
Db 1028 AGAGTCTCTGAGGATGATGAGAAATGGG-----AGATTCCCACTGTTATAGCTGCT 1078
Qy 164 SerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGlu 183
Db 1079 GGTCA-----AGTCGAGAAATTCAGTTGGAGGAGGTTATGTTAATGGANA 1126
Qy 184 ValSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySerAla 203
Db 1127 -----CATGGGCTTCATAAGCGTGTGCACCATATCCATCATCTGAAGCTGGTGAAGTCT 1180

Qy 204 LysTrpAsp-----GluLysLysGluValSer----- 212
Db 1181 CA-TGGAAATGTTAACTTACATATAGATTTAAAGAAATGCTCTCACAGTGAATGATTAGTAGG 1239
Qy 212 ----- 212
Db 1240 GTCATGTCATATCTCCATATGTGCAAAATAACATAAGTATGAGGCTTCCAGCTTAAATAGTA 1299
Qy 212 ----- 212
Db 1300 GATAGGACATAGTTTTCATAAACATGGACTTTGGGTTCTATTACATTTCTTCTATGAAAT 1359
Qy 212 ----- 212
Db 1360 TCATCAGCAGACCTTTTCTTAACAATACTTCTCTTTTGTGTTATGTGCTAAATTTTAAAT 1419
Qy 213 -----TrpLysGluArgMetAspAspTrpLysSerLysGlnGly 225
Db 1420 GTGTAGGAGTAGTGGGAGGATGCGGGAAGAAATGCGATGACTGGAAGCTCCAGCATGGA 1479
Qy 226 IleLeuGlyGlyGlyAlaAspProGlu---AspMet-AspAlaAspValAlaLeuAsn-- 243
Db 1480 AATCTTGGGCGAGAACCCAGATGATGATCCTGAGATGGGACTGTAATGCCCTCCACAAACAT 1539
Qy 243 ----- 243
Db 1540 TTATCTAAGACATCAGTTTTTGTATGATTGATGCTTACAAATTTTGGATTGAC 1599
Qy 244 -----AspGluAlaArgGlnProLeuSerArgLysValSerIleAl 257
Db 1600 TGGAAATGGCTGTAGGATCGAGGACGCGACCACTCTCGCGGAAGTTCCCATGTC 1659
Qy 257 aSerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPh 277
Db 1660 CTCAAGCAAGATCAATCCATCGATCGATGGTCACTGTTAGCTTGTGATTTCTAGCAGT 1719
Qy 277 ePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrpLeuValSe 297
Db 1720 TTTTTCGGTATAGGCTTTGAATCCAGTCATGATGCTCTGCGATTATGGCTGACCTC 1779
Qy 297 rIleIleCysGluIleTrpPheAlaIleSerTrpIleLeuAspGlnPheProLysTrpPh 317
Db 1780 TGTGATCTGTGAATCTGTTGCTGCTCTCTGATGATCTTGAATCAGTTCCTCCCAAGTGGTT 1839
Qy 317 eProIleAspArgGluThrTyrLeuAspArgLeuSerLeu----- 330
Db 1840 CCCTATTGAACGTGAGACCTATCTAGATCGGCTTCCCTCAGGTAAATCCACAGATTCT 1899
Qy 330 ----- 330
Db 1900 CAAGTAGAAGTCTTAAATCTATGACGTTGGAGTTTGGATGATAAATATTTTGTTTATA 1959
Qy 331 -ArgTyrGluArgGluGlyGluProSerLeuLeuSerAlaValAspLeuPheValSerTh 350
Db 1960 CAGGTACGAGAGAGAGGTGAACCAATATGCTTGGCCCTCTAGATGCTCTTTGTGAGTAC 2019
Qy 350 rValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAl 370
Db 2020 GGTGACCCCATGTAAGGAGGCTCCCTCGTCGTACATCAACACACTGTGCTGTCAATCTGGC 2079
Qy 370 aValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLe 390
Db 2080 CATGACTTACCAGTTGAGAAATCTCTGTATGCTCTCTGACGCGGTCTTCAATGCT 2139
Qy 390 uThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrpValProPheCysLys 410
Db 2140 TACATTCGAATCTCTCTCGGAAATCTGCTGAGTTTGCAGAAAAATGGGTTCCCTCTGTAA 2199
Qy 410 sLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrLe 430
Db 2200 GAAATTTCCATAGAGCCAGGACCCGAGATGTACTTTCAGTTGAAGTTGATTACT 2259
Qy 430 uLysAspLysValGlnProThrPheValGlnGluArgAlaMetLys----- 446

```
Db 2260 CAGGCAAGTCCACCCACATTTGTTAAGAACGTCGAGCCATGAAGGTCAAGTGTATA 2319
QY 446 ----- 446
Db 2320 TCACCTGATCTAGTTATACCACCACCACCTCTTTTCACTATAATCTAAACTTCAATAGTGA 2379
QY 447 -----ArgGluTyrGluGluPheValArgIleAsnAlaLeuValAl 461
Db 2380 AATTGACATATACAGAGAGAGATATGAGGAGTTCAGGTTCAGGATCAACGCTCAAGTGC 2439
QY 461 alyAsnAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProG 481
Db 2440 GAAGGCTCAAGGTTCTCTAGAGTTGGATTCAGATGGAACCGTGGCCAGG 2499
QY 481 yAsnAlaMetLysValProAlaGluGlyTrpIleMetLysAspGlyThrProTrpProG 501
Db 2500 GAACAACACCAAGGACCCCGGTATGATCCAAAGTCTTCTCGGCCACAGCGAGGATT 2559
QY 501 sAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProG 521
Db 2560 TCATGTCGAAGGCGATGAGCTTCTCGGCTTGTGTACGTGTCCCGTGAAGCGTCTGG 2619
QY 521 yPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuIle----- 535
Db 2620 TTTTCAACACCAAGAGCTGGGCCATGAATGCCCTGGT-AAATTTCTTGATCTGCC 2678
QY 535 ----- 535
Db 2679 TTGGACCAACAAGAACTGATCTCGTGTCTTGGACCTAACCTGATCTCTGTGAGGT 2738
QY 536 -ArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuLeuAspCysAspHisT 555
Db 2739 TCGAGTCCAGCGCGTACTCACAAATGCTCTTTCATGCTGAACTTGGACTGTGATCACTA 2798
QY 555 rIleAsnAlaSerLysAlaIleArgGluAlaMetCysPheLeuMetAspProGluValG 575
Db 2799 TGTAACCAACAGCAGCGCGTGGAGGAAGCAATGTGTTTTGTGGATCCTCAGATTGG 2858
QY 575 yArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspAr 595
Db 2859 AAAGAAGGTCTCTATGTTCACTTCCCTCAAGGTTTGATGGCATTGACACAAACGATCG 2918
QY 595 gTyrAlaAsnArgAsnThrValPhePheAsp----- 605
Db 2919 TTACGCCCAACAGAAACACAGTCTTCTTGTGATGAAGACTCAATTCATATTTTCCAACTT 2978
QY 606 -----IleAsnMetL 609
Db 2979 CTGGTCTAAATGAANAATGACCTCTTGCTTACACTCTTGTGTTGCTACAGATCAATATGA 3038
QY 609 ysGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAspArg 629
Db 3039 AAGTCTAGATGAATCCAGGTCAGTTTACGTTGTTGTTGTTGTTGTTTCAACGAC 3098
QY 629 InAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysA 649
Db 3099 AAGCTCTGTATGTTATGAACCAACAAAGGCTCTAAACGTCCTCAAGATGATAAGCTGTG 3158
QY 649 sPCysCysProCysPheGly-----ArgLysLysArgGlyHisAlaLysAspGlyLeu- 666
Db 3159 GTTGTTGCTTGTCTTGGCCCGGAGAAAGAAATTAAGAAATTTTCCAAAGATGATCATGA 3218
QY 667 -----ProG 668
Db 3219 ATGGTGACGTAGCAGCCCTTGGAGGTAATATATCCCAACACCTTATATATATCAGTCCAT 3278
QY 668 luGlyThrAlaAspMet-----G 674
Db 3279 TCCTT-GCAGTAGATTTGCTTTATGTTGAAATCTTCGGGATCTGATGATGTTTTTGGCAG 3337
QY 674 lyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyG 694
Db 3338 GAGCAGAAAGGTGATTAAGGAGCATTTGATGTCGAAATGAACCTTTGAGAAAAACATTTGGGC 3397
QY 694 InSerAlaAlaPheValThrSerThrLeuMetGluGluGlyValProProSerSerS 714
Db 3398 AATCATCCATCTTTTAACCTCAACTTTGATGTAAGAGAGGTGGTGTCTCTCGTCATCAA 3457
QY 714 erProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysT 734
Db 3458 GTCTCTGCAAGTGTCTCTTAAGAGGCAATCCATGTATAAGCTGCGGTATGAAGACAAGA 3517
QY 734 hrAspTrpGly----- 737
Db 3518 CTGAGTGGGAACTAGGCTAATAATACTGAATCGTAGAAATCACCTCTTATTGTGTGATT 3577
QY 738 -----LeuGluLeuGlyTTrpIleT 744
Db 3578 TAGTAGCTGGCCATTTGTAATAAAACGTTTGTGTATCTCGAAATTCAGCTGGGTGGATCT 3637
QY 744 yzGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTTrpArgS 764
Db 3638 ATGGCTCTATCACAGAGGATATTTTGACGGGATTCAGATGTCATTCGCGTGGATGGAGT 3697
QY 764 erValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerA 784
Db 3698 CTATTTACTGATGCTTAAGAGGCTGCTATTCAAAGGTTTCAGCTCTTATTAATCTATCAG 3757
QY 784 sPArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhePheSerArgH 804
Db 3758 ACAGGTTAAACCCAGGTTTGGTGGGCACITGGATCGGTTGAGATATTTTTCAGCCGCG 3817
QY 804 isSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaT 824
Db 3818 ACAGTCTCTCTGCTGATGGCTACAAAGAGGACAACTCAAGTGGCTTTCAGCGTTCCTT 3877
QY 824 yzIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuP 844
Db 3878 ATGCCACACAAACAATCTACCCCTTCATCTATACCACTCTTTCGCTTACTGTATCTTTC 3937
QY 844 roAlaValCysLeuLeuThrGlyLysPheIleMetProSer----- 857
Db 3938 CAGCCATCTGCTCTCTTACTGACAAATTCATCATGCCACC-GGTTAGTAAATATATCAGA 3996
QY 857 ----- 857
Db 3997 GAAAGACACTTAGAAGCTGCATCAAAATGTGCTAACTATCTGTTTCCCAATTTTCTTTC 4056
QY 858 --IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrG 877
Db 4057 AGATAAGCACATTTGCTAGTCTCTTCTTCATCTCACTGTTTATGTCGATCATTTGTAACGG 4116
QY 877 lyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTyrTrpArgAsnGluGlnP 897
Db 4117 GAATCTTGGATTCAGATGGAGCGGAGTTAGCATTTGAGAAATGGTGGAGAACGAGCAAT 4176
QY 897 heTrpValIleGlyGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysV 917
Db 4177 TCTGGCTCATTTGGAGGAATCTCAGCTCATCTCTTTCGCGTGTGCAAGGCTCTCTCAAAA 4236
QY 917 alLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspA 937
Db 4237 TCTTAGCAGGATTTGACAAACTTCACCGTCATCAAGGCAACA-----GATGATG 4290
QY 937 spGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProThrThrL 957
Db 4291 ATGACTTTGGAGAACTTTAGCATTCAATATGGACACACACTGCTGATCCCTCCACAACTG 4350
QY 957 euLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnGlyT 977
Db 4351 TCTTAATCATAAACATTTGTCGCGTGTGTCAGGCACTCTCAGATGCCATTAAACAATGGAT 4410
QY 977 yzGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisL 997
Db 4411 ATCAGTCTTGGGACCTCTATTTGGTAAACTCTTCTTCTCTCTTCTTCTTCTTCTTCTTCT 4470
```

Qy 997 eutyPrProPheLeuLeuGlyLeuMetGlyArgGlnAsnArgThrProThrValValVal 1017
 Db 4471 TCTACCCATTCTCTCAAGGCTGATGGGTAGACAGAAACAGACACCAACCACTTGTGTGA 4530
 Qy 1017 leTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTTPValArgIleAspProPheI 1037
 Db 4531 TTTGGTCAGTGTATTGGCATCTATCTCTCTTCTTGGTGGTAAGAATTGATCCCTTTG 4590

Qy 1037 leValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
 Db 4591 TGCTCAAGACCAAGGACCTGACATCTCAAGTGTGGCATCACTGC 4637

RESULT 15
 AF091713 7234 bp DNA linear PLN 27-JUN-1999
 LOCUS Arabidopsis thaliana cellulose synthase catalytic subunit (IRX3)
 DEFINITION gene, complete cds.

ACCESSION AF091713 GI:5230422

VERSION AF091713.1

KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 7234)
 AUTHORS Taylor,N.G., Scheible,W.R., Cutler,S., Somerville,C.R. and Turner,S.R.

TITLE The irregular xylem3 locus of Arabidopsis encodes a cellulose synthase required for secondary cell wall synthesis

JOURNAL Plant Cell 11 (5), 769-780 (1999)

MEDLINE 99264300

PUBMED 10330464

REFERENCE 2 (bases 1 to 7234)

AUTHORS Taylor,N.G. and Turner,S.R.

TITLE Direct Submission

JOURNAL Submitted (09-SEP-1998) Biological Sciences, 3.614 Stopford Building, University Of Manchester, Oxford Road, Manchester M13 9PT, UK

FEATURES

source

1. 7234

/organism="Arabidopsis thaliana"

/mol_type="Genomic DNA"

/cultivar="Landsberg erecta"

/db_xref="taxon:3702"

/chromosome="v"

<92..>4640

/gene="IRX3"

/note="irregular xylem3"

join(<92..166,569..764,894..1172,1427..1520,1615..1881,1963..2308,2396..2659,2737..2949,3029..3242,3338..3534,3625..3978,4059..>4640)

/gene="IRX3"

/product="cellulose synthase catalytic subunit"

join(92..166,569..764,894..1172,1427..1520,1615..1881,1963..2308,2396..2659,2737..2949,3029..3242,3338..3534,3625..3978,4059..>4640)

/gene="IRX3"

/codon_start=1

/product="cellulose synthase catalytic subunit"

/protein_id="AAD40885.1"

/db_xref="GI:5230423"

/translation="MEASGLVAGSHRNELVVIHNEEPKPLKNLDQCFCEICDQI

GLTVRGLDFVACNECGFPACRCPCEYERREGTQNCQCKTRYKRLGSPRVEGDEE

DLDDIEYEINIEHQDKEKHAEMLYGKMSYGRGEDDENGRRFPVPLAGHSGFFV

GGYNGENHGLHVRVHPVPSAGSGGWRERMDWLQHGNIIGPEDDPEMGLIDE

AKQPLSRKVPASSKINPYRMVIVARLVILAVFLRYLLNPVHDALGLWLSVCEIW

FAVSWLDQFPKPFIERETLYDLRLSLRYEREENMLAPVDVFSVTDVPLKEPLVT

SNTVLSILAMDYPVEKISCTYSDDGASMLTFESSETAFARKWVPFCCKFSIEPRAP

ENYFTLVLDYLDQKVHPTFVKERRAMKEVEFKVIRINQAVAKASVPLEGIMQDGT

PWPGNTKDHQPMI QVFLHSGGFDVEGHELPLVYVSREKRPQGHKKAGAMNALV

RVAGVLITWAPFNLNDCDHYVNSKAVREAMCFIAMDPIQKVKVCVQFPQFDGIDTN

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 7234
 Score: 4224.50 Matches: 856
 Percent Similarity: 61.35% Conservative: 87
 Best Local Similarity: 55.69% Mismatches: 90
 Query Match: 74.80% Indels: 508
 DB: 8 Gaps: 20

US-10-627-132-30 (1-1052) x AF091713 (1-7234)

Qy 1 MetGluAlaSerAlaGlyLeuValAlaGlySerHisAsnArgAsnGluLeuValLeu 20
 Db 92 ATGGAAGCTAGCCCGCTCTTGTCCGCGGTCTCTATAACCGTAATGAAGCTAGTCGTCATT 151
 Qy 21 ArgGlyHisGluAsp----- 25
 Db 152 CACAAACCATGAGAGGTTTTTCACATTACTTTTTCTCATCACTTACCAAGTTTTTTTT 211
 Qy 25 ----- 25
 Db 212 TACCACGCTAGTAAATAATTATTCATTTTTCGTTTTTATTTGGTTACTATTTCTTAAA 271
 Qy 25 ----- 25
 Db 272 AATTGCTATGTTTGGGAATAAAATATGATCATTTTTTTTAAATCATCTCTTTATTATGA 331
 Qy 25 ----- 25
 Db 332 GACAAAAATTATAATCTGTATCTGTAGTTGCAATAATGTTGAGAAAAATTCATATCTT 391
 Qy 25 ----- 25
 Db 392 TGTTAGCAACATAATAATTTTGTGTAATAATTAAGTTGAGAGTCAGGTTTAAACATT 451
 Qy 25 ----- 25
 Db 452 TTAATCGCTGTCATTTTTTTTATTTATCTTTACTTATAGAACTTTTTTTTGTGATTAG 511
 Qy 26 -----Pro 26
 Db 512 AATTTTCAGGTTTTTATTTTCGTTCTTTAAGGAACTTTAAACTTTTGTGTAATTATAACAGCCA 571
 Qy 27 LysProLeuArgAlaLeuSerGlyGlnValCysGluLeuCysGlyAspGluValGlyLeu 46
 Db 572 AAGCCTCTGAAGAATCTAGATGGCAATTTCTGTGAGATATGTGGAGATCAATTTGGTTA 631
 Qy 47 ThrValAspGlyAspLeuPheValAlaCysAsnGluCysGlyPheProValCysArgPro 66
 Db 632 ACAGTAGAAGGAGACCTCTTCGTAGCTTGCATGAGTGTGTTTTCCGGCGGTAGACCT 691
 Qy 67 CysTyrGluTyrGluArgArgGluGlyThrGlnAsnCysProGlnCysLysThrArgTyr 86
 Db 692 TGCTATGATGACGAG 751
 Qy 87 LysArgLeuLys----- 90
 Db 752 AAGCGTCTCAGAGGTAAGTTATTTATTAATCTCCCTCTGCTCTTTGTGTTGTCGACGAAA 811
 Qy 90 ----- 90
 Db 812 TGCCTCTATGAAAAATTTAAAAAGGCTGTTCTCTTTTATTTAGTTGAACTTGAGAGTAATG 871
 Qy 91 -----GlySerProArgValAlaGlyAspAspGluGluAsp 103

DRYANRNTVFFDINMKGLDGIQGPVVVGTGCVFKRQALYGYBPPKGPKPKNMISCCCC
 PCFRRRNKFKSKNDMNGDVAALGAEKDEKHELMSENNFEKTFGQSSIFVSTLMEE
 GGVPPSSPAVLKKEALHVI SCGYEDKTEWGTGLWYIGSITEDIUTGFKMHCRCWRSS
 IYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEIFPSRHSPLWYGVKGLKWLRRF
 AVANTTIYPTSIPLLAYCILPAICLLTDKFIIMPFIISTFASLFFISLKSIIITGLIE
 LRWSGVSIENWRNEQFVWIGGISAHLFVAVOGLLKL LAGIDTNFTVTSKATDDDDFG
 ELYAFKWTLLIIPPTVLINIVGVVAGISDAINNGYSQSGPLFGKLFPSPWVILHY
 PFLKGLMGRQNRPTTIVWISVLLASIFSLWVRIDPFLVLTGKGPDTSCGKNC"

Db 872 ATCTGTTTTTGGTTCTCAAGAGCCCAAGAGTGGAGGAGATGAAGACCAAGAGAT 931
Qy 104 IleAspAspLeuGluHisGluPheAsnIleAspAspGluAsnGlnArgGlnLeuGlu 123
Db 932 ATTGATGATATTGAGTATGAATTTAATATCGAACATGAACAAGAT----- 976
Qy 124 GlyAsnMetGlnAsnSerGlnIleThrGluAlaMetLeuHisGlyArgMetSerTyrGly 143
Db 977 -----AAGCATGAAGCATTTCTGCTGAGGCTATGCTTTATGGGAAATGAGCTATGGA 1027
Qy 144 ArgGlyProAspAspGlyAspGlyAsnAsnThrProGlnIleProProIleIleThrGly 163
Db 1028 AAGAGTCTGAGATGATGAGATGG-----AGATTCACCTCTGTTATAGTGTGT 1078
Qy 164 SerArgSerValProValSerGlyGluPheProIleThrAsnGlyTyrGlyHisGlyGlu 183
Db 1079 GGTGAT-----AGTGAGAAATTTCCAGTTGGAGGAGTTATGTTAATGGAGAA 1126
Qy 184 ValSerSerLeuHisLysArgIleHisProTyrProValSerGluProGlySerAla 203
Db 1127 -----CATGGCTTTCAATGAAGGTGTGCACCCATATCCATCATCTGAAGCTGGTGTCT 1180
Qy 204 LysTyrAsp-----GluLysLysGluValSer----- 212
Db 1181 CATGGAATGTTAACTTACATATAGATTTAAGATGTTCTACAGTATGATTAGTTAGG 1239
Qy 212 ----- 212
Db 1240 GTCATGCATATCTCCATATGTGCAATAACATAAGTATGAGGCTTCAGCTTAATAGTA 1299
Qy 212 ----- 212
Db 1300 GATAGGACATAGTTTCAATAACATGAGCTTTGGGTTCTATTACATTTCTTCTATGAAAT 1359
Qy 212 ----- 212
Db 1360 TCATCAGCAGACCTTTTCTAACAAATCTTTCTTTTGTGTTATGTTGTTAATTTAAT 1419
Qy 213 -----TrrLysGluArgMetAspAspTrrLysSerLysGlnGly 225
Db 1420 GTGGTAGGAGTGAGGAGGATGGCGGNAAGAAATGGATGACTGGAGCTCCAGCATGA 1479
Qy 226 IleLeuGlyGlyAlaAspProGlu---AspMet-AspAlaAspValAlaLeuAsn-- 243
Db 1480 AATCTTGGGCCAGAACCAAGATGATCTCTGAGATGGGACTGTAATGCTCCCAACAACAT 1539
Qy 243 ----- 243
Db 1540 TTATCTAAGACATCAGTTTGTATGATTTGGATTCATGCTTACAAAAATTTTGGATTTGAC 1599
Qy 244 -----AspGluAlaArgGlnProLeuSerArgLysValSerIleAl 257
Db 1600 TGGAAATGGCTGTAGATCGACGAGCGACGAGCCACTCTCGCGAAAGTTCCCATTTGC 1659
Qy 257 sSerSerLysValAsnProTyrArgMetValIleValValArgLeuValValLeuAlaPh 277
Db 1660 CTCGAAGCAAGATCAATCCATATCGGATGTCATGCTTGTAGGCTTGTGATTTAGCAGT 1719
Qy 277 ePheLeuArgTyrArgIleLeuHisProValProAspAlaIleGlyLeuTrrLeuValSe 297
Db 1720 TTTTCTGCGGTATAGGCTCTTGAATCCAGTGCATGATGCTCTGGGATTTATGGCTGACCTC 1779
Qy 297 rIleIleCysGluIleTrrPheAlaIleSerTrrIleLeuAspGlnPheProLysTrrPh 317
Db 1780 TGTGATCTGTGAATCTGGTTCGCTGCTCTCTTGGATTTCTTGATCAGTTCCTCCCAAGTGT 1839
Qy 317 eProfileArgGluThrTyrLeuAspArgLeuSerLeu----- 330
Db 1840 CCTATTGAACGTGAGACCTATCTAGATCGGCTTTCCCTCAGGTAAAAATCCACAGATTCT 1899
Qy 330 ----- 330

Db 1900 CAACTAGAAAGTCTTAAAAATCTATGACGTTGGAGTTTGGATGAAATATTTTTTGTATTATA 1959
Qy 331 -ArgTyrGluArgGluGlyGluProSerSerLeuLeuSerAlaValAspLeuPheValSerPh 350
Db 1960 CAGGTACGAGAGAGAGAGGTGAACCAATATGCTTCCCTCTGTAGATGCTTTTGTCAGTAC 2019
Qy 350 rValAspProLeuLysGluProProLeuValThrAlaAsnThrValLeuSerIleLeuAl 370
Db 2020 GGTGACCACTTGAAGAGGAGCTCCCTCGTCACATCCAAACATGCTGCTCAATCTTGC 2079
Qy 370 aValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLe 390
Db 2080 CATGACATCCAGTGTGAGAAATCTCTGCTATGCTCTGACGAGGTGCTTCAATGCT 2139
Qy 390 uThrPheGluSerLeuSerGluThrAlaGluPheAlaArgLysTrrPValProPheCysLy 410
Db 2140 TACATTCGAATCTCTCTCGAAACTGCTGAGTTTCAAGAAATCGGTTCCCTTCTGTAA 2199
Qy 410 sLysPheGlyIleGluProArgAlaProGluPheTyrPheSerLeuLysValAspTyrIle 430
Db 2200 GAAATTTCTCCATAGAGCCACGCGACCGAGATGTACTTCAAGTTGAAAGTTGATTATCT 2259
Qy 430 uLysAspLysValGlnProThrPheValGlnGluArgAlaMetLys----- 446
Db 2260 CAGGACAAAGTCCACCAACATTTGTTAAGGAAGCTGAGCCATGAAGGTCAAGTGTATA 2319
Qy 446 ----- 446
Db 2320 TCACCTGATCTAGTTATATACCAACCCCATCTTTCACTATATCTTAACTTCATAGTGA 2379
Qy 447 -----ArgGluTyrGluGluPheLysValArgIleAsnAlaLeuValAl 461
Db 2380 AATTGACATATCAGAGAGAGATATGAGGAGTCAAGGTCAGGATCAACGCTCAAGTGC 2439
Qy 461 aLysAlaMetLysValProAlaGluGlyTrrIleMetLysAspGlyThrProTrrProGl 481
Db 2440 GAAGCCCTCAAGGTTCTCTAGAGGTTGATCATGCAAGATGAAACACCGTGCCAGG 2499
Qy 481 yAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyHi 501
Db 2500 GAACAACACCAAGGACCAACCCCGTATGATCCAAAGTCTTCTCGGCCACACGCGAGGATT 2559
Qy 501 sAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGl 521
Db 2560 TGATGTGCAAGGCGATGAGCTTCTCGGCTTGTGTACGTGTCCGCTGAGAAAGCTGCTGG 2619
Qy 521 yPheGlnHisHisLysLysValAlaGlyAlaMetAsnAlaLeuIle----- 535
Db 2620 TTTTCAACACCAAGAAAGCTGGCGCATGAATGCCCTGGT-AAATTTCTTGTATCTGCC 2678
Qy 535 ----- 535
Db 2679 TTGGACCAACAAGAAATGATCTCCGTGCTTGGACCTTAACCTGATACTTCTGTCAAGT 2738
Qy 536 -ArgValSerAlaValLeuThrAsnAlaProPheMetLeuAsnLeuAspCysAspHisTy 555
Db 2739 TCGATGTGCGAGCGGTACTCACAATGCTCTTTCATGCTGAATCTGGACTGTGATCACTA 2798
Qy 555 rIleAsnAsnSerLysValIleArgGluAlaMetCysPheLeuMetAspProGlnValGl 575
Db 2799 TGTAAACAACAGCAAGCGCGTGGAGGAAGCAATGTGTTTTTGTGGATCTCTCAGATGG 2858
Qy 575 yArgLysValCysTyrValGlnPheProGlnArgPheAspGlyIleAspValHisAspAr 595
Db 2859 AAAGAGGTCTGCTATGTTTCAGTTCCTCAAGGTTTGTATGCGATTTGACACAACGATCG 2918
Qy 595 gTyrAlaAsnArgAsnThrValPhePheAsp----- 605
Db 2919 TTACGCCCAACAGAAACACAGTCTTCTTTGTATGTAAGACTCAATTCTATTTTCCAACTT 2978
Qy 606 -----IleAsnMetL 609
Db 2979 CTGGTCTAAATGAAATGTACCCCTCTTGTCTTACACTCTTGTGTTGCTACAGATCAATATGA 3038

```
QY 609 ySgLyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheArgG 629
Db 3039 AAGGCTAGATGGAAATCAAGTCCAGTTTACGTTGGTACTGGTTGGTTTCAAGCAC 3098
QY 629 lnAlaLeuTyrGlyTyrAsnProProLysGlyProLysArgProLysMetValThrCysA 649
Db 3099 AAGCTCTGTATGGTTATGAACCAACCAAGGGTCTTAACGTCCTCAAGATGATAAGCTGTG 3158
QY 649 spCysCysProCysPheGly-----ArgLysLysArgLysHisAlaLysAspGlyLeu- 666
Db 3159 GTTGTCTCTCTTCTTGGCGCGGAGAAAGAAATAAGAAATTTTCCAAAGATGACATGA 3218
QY 667 -----ProG 668
Db 3219 ATGGTGAGGTAGAGCCCTGGAGGTAATAATATCCCAACCAACCTTATAATATCATGCTCAT 3278
QY 668 luGlyThrAlaAspMet-----G 674
Db 3279 TCCTT-GCAGTAGATTTCTGTTTATGTTGGAATCTTTGCGGATCTGATAGTGTGTTTGGCAG 3337
QY 674 lyValAspSerAspLysGluMetLeuMetSerHisMetAsnPheGluLysArgPheGlyG 694
Db 3338 GAGCAGAAGGTGATAAGGAGCATTTGATGTCGAAATGAACCTTTGAGAAAAACATTTGGGC 3397
QY 694 lnSerAlaAlaPheValThrSerThrLeuMetGluGlyGlyValProProSerSers 714
Db 3398 AATCATCATCTTTGTAACTCAACTTTGATGGAAGAGGGTGTTCCTCCGTCATCAA 3457
QY 714 erProAlaAlaLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysT 734
Db 3458 GTCTCGAGTCTCTCTTAAAGAGGCAATCCATGTCATAAGCTGCGGTTATGAAGACAAGA 3517
QY 734 hrAspTrpGly----- 737
Db 3518 CTGAGTGGGAACTGAGGTAAATACTGAATCGTAGAAATCACCTCTTATTATTGTGATT 3577
QY 738 -----LeuGluLeuGlyTrpIleT 744
Db 3578 TAGTAGTGGCCATTGTAAAAAAGCTTTGTGTATCTCGAAATTCAGCTGGTTGGATCT 3637
QY 744 yrGlySerIleThrGluAspIleLeuThrGlyPheLysMetHisCysArgGlyTrpArgS 764
Db 3638 ATGGCTCTATCACAGAGGATATTTGACGGGATTCAGATGCATTCGCCGTGATGGAGGT 3697
QY 764 erValTyrCysMetProLysArgAlaAlaPheLysGlySerAlaProIleAsnLeuSerA 784
Db 3698 CTATTTACTGTCATGCCTAAGAGGCCGTCATTCAAAGGTTTCAGCTCTCTATTAAATCTATCAG 3757
QY 784 spArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIlePhePheSerArgH 804
Db 3758 ACAGGTTAAACAGGTTTTGGCTTGGGCACCTTGGATCGGTTGAGATATTTTTCAGCCGCG 3817
QY 804 isSerProLeuLeuTyrGlyTyrLysAsnGlyAsnLeuLysTrpLeuGluArgPheAlaT 824
Db 3818 ACAGTCTCTCTGGTATGGCTACAAAGGAGGCAAACTCAAGTGGCTTCAGCGTTTTGCTT 3877
QY 824 yrIleAsnThrThrIleTyrProPheThrSerLeuProLeuLeuAlaTyrCysThrLeuP 844
Db 3878 ATGCCAACACACAACATCTACCCCTTCACATCTATACCACATCTTCTTGGCTACTGTATCCTTC 3937
QY 844 roAlaValCysLeuLeuThrGlyLysPheIleMetProSer----- 857
Db 3938 CAGCCATCTGCTCTCTTACTGACAAATTCATCATGCCACC-GGTTAGTAAAAATATATCAGA 3996
QY 857 ----- 857
Db 3997 GAABAGCACTTAGAAGTCATCAAAATGTGTAATCTATCTGTTTTTCCCAATTTTCTTTC 4056
QY 858 ---IleSerThrPheAlaSerLeuPhePheIleAlaLeuPheMetSerIlePheAlaThrG 877
Db 4057 AGATAAGCACATTTGTAGTCTCTTCTCATCTCTCTGTTATGTCGATCATCTGTAACGG 4116
```

```
QY 877 lyIleLeuGluMetArgTrpSerGlyValSerIleGluGluTrpTrpArgAsnGluGlnP 897
Db 4117 GAATCTTGGAAATTGAGATGCGGAGTTAGCATTTGAGGAATGCTGGAGAAACGAGCAAT 4176
QY 897 heTrpValIleGlyValSerAlaHisLeuPheAlaValValGlnGlyLeuLeuLysV 917
Db 4177 TCTGGGTCATTTGGAGGAATCTCAGCTCATCTCTTTCGGTGTGTCGAAGGTCTCTCAAAA 4236
QY 917 alLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaThrGlyAspGluAspA 937
Db 4237 TCTTAGCAGCATTTGACAAACTTCACCGTCAATCAAAAGGCAACA-----GATGATG 4290
QY 937 spGluPheAlaGluLeuTyrAlaPheLysTrpThrThrLeuLeuIleProProThrThrL 957
Db 4291 ATGACTTTGGAGAACTTTAGGCATTCAAATGGCAACACTGCTGATCCCTCCCAACAACGTG 4350
QY 957 euLeuIleIleAsnValIleGlyValValAlaGlyIleSerAspAlaIleAsnAsnGlyT 977
Db 4351 TCTTAATCATAAACATTTGTCGGCTTGTTCAGGCATCTCAGATGCCATTTAAACAATGGAT 4410
QY 977 yrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisL 997
Db 4411 ATCAGTCTTGGGACCTCTATTTGGTAAACTCTTCTTCTCTTTGGGTCAATTGTTCAATC 4470
QY 997 euTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrValValValI 1017
Db 4471 TCTACCCATTTCTCTCAAGGTCTGATGGTAGACAGAACACACCAACCATTTGTGTGA 4530
QY 1017 leTrpSerIleLeuLeuAlaSerIlePheSerLeuLeuTrpValArgIleAspPropheI 1037
Db 4531 TTTGGTCAGTGTATTGGCATCTATCTCTTTTGTCTTTGGGTAAAGAAATGTATCTTTTG 4590
QY 1037 leValArgThrLysGlyProAspValArgGlnCysGlyIleAsnCys 1052
Db 4591 TGCTCAGACCAAGGACCTGACACTTCCAAAGTGTGGCATCAACTGC 4637
```

Search completed: March 12, 2005, 13:08:15

Job time : 7860 secs